

11/5,K/1 (Item 1 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2008 The Thomson Corporation. All rts. reserv.

0017320652 - Drawing available
WPI ACC NO: 2008-B41092/200810
XRPX Acc No: N2008-110251

Annotation decoding

Patent Assignee: DING J (DING-I)

Inventor: DING J

Patent Family (1 patents, 1 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update
CN 101038585	A	20070919	CN 200610070821	A	20060313	200810 B

Priority Applications (no., kind, date): CN 200610070821 A 20060313

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing	Notes
CN 101038585	A	ZH		1		

CN A
NOVELTY - Annotation decoding universally labeled or substituted from letters of different languages of countries all over the world is a simple, practical, understandable and universally accepted digital transforming code with which language characters or letters of countries all over the world can be substituted and different language letters of characters of countries all over the world can be universally labeled and expressed. It is comprised of ten Arabic number identifiers, arranged digital code in alphabetic order, Chinese, combined character, spelling rule and phonetic sign. If language character / letter of each country can set up letter, term, sentence library by the use of the code, which is favor for the digital code of the language letter / character of each country, the letter, term, sentence can be found with the help of the voice of the annotation decoding and can be directly transformed by the use of the computer. Therefore, the invention is an aide for international exchange.

Title Terms/Index Terms/Additional Words: DECODE

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0017/28	A	I	F	20060101
G06F-0017/28	C	I		20060101

File Segment: EPI;

DWPI Class: T01

Manual Codes (EPI/S-X): T01-D02; T01-J14

...the code, which is favor for the digital code of the language letter / character of each country, the letter, term, sentence can be found with the help of the voice of the annotation decoding and can be directly transformed...

Original Publication Data by Authority

Original Abstracts:

...the code, which is favor for the digital code of the language letter / character of each country, the letter, term, sentence can be found with the help of the voice of the annotation decoding and can be directly transformed...

11/5,K/2 (Item 2 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2008 The Thomson Corporation. All rts. reserv.

0012799914
WPI ACC NO: 2002-656541/200270
Related WPI Acc No: 1999-404693; 2001-601873; 2006-687590
XRPX Acc No: N2002-518998

Computer implemented document summarizing method for authors and readers, involves constructing and inserting sentence based summary of document's writings at beginning of document

Patent Assignee: COKUS S J (COKU-I); DOLAN W B (DOLA-I); FEIN R A (FEIN-I); FRIES E J (FRIE-I); MESSERLY J (MESS-I); MICROSOFT CORP (MICT); THORPE C A (THOR-I)

Inventor: COKUS S J; DOLAN W B; FEIN R A; FRIES E J; MESSERLY J; THORPE C A

Patent Family (2 patents, 1 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update
US 20020103836	A1	20020801	US 1999289085	A	19990408	200270 B
			US 200274951	A	20020211	
US 7051024	B2	20060523	US 1999289085	A	19990408	200635 E
			US 200274951	A	20020211	

Priority Applications (no., kind, date): US 1999289085 A 19990408; US 200274951 A 20020211

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 20020103836	A1	EN	10	3	Continuation of application US 1999289085
US 7051024	B2	EN			Continuation of application US 1999289085
					Continuation of patent US 6349316

Alerting Abstract US A1
NOVELTY - A sentence based summary of a document's writings is constructed and inserted at the beginning of the document.
DESCRIPTION - INDEPENDENT CLAIMS are included for the following:

- 1.Word processing application;
- 2.Electronic mail application;
- 3.Internet web browser application;
- 4.Computer summarizing document; and
- 5.Document file.

USE - For summarizing documents helpful in assisting authors and readers.
ADVANTAGE - Enables authors to automatically create summaries of their writings with improved quality, in a convenient and useful way to the author.

Title Terms/Index Terms/Additional words: COMPUTER; IMPLEMENT; DOCUMENT; SUMMARY; METHOD; READ; CONSTRUCTION; INSERT; SENTENCE; BASED; BEGIN

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0017/00	A	I	F	B	20060101
G06F-0017/27	A	I		R	20060101
G06F-0017/30	A	I	L	B	20060101
G06F-0017/30	A	I		R	20060101
G06F-0017/00	C	I	L	B	20060101
G06F-0017/27	C	I		R	20060101
G06F-0017/30	C	I	L	B	20060101
G06F-0017/30	C	I		R	20060101

US Classification, Issued: 707531, 7076, 715500, 715531

File Segment: EPI;

DWPI Class: T01

Manual Codes (EPI/S-X): T01-E01B; T01-J11A1; T01-N01C; T01-S02

...authors and readers, involves constructing and inserting sentence based summary of document's writings at beginning of document

...A sentence based summary of a document's writings is constructed and inserted at the beginning of the document.

Original Publication Data by Authority

Original Abstracts:

...Summarizer performs a statistical analysis to generate a list of ranked sentences for consideration in the summary. The summarizer counts how frequently content words appear in a document and produces a table correlating the content words with their corresponding frequency counts. Phrase compression techniques...

...inclusion of the sentence have been satisfied. The summarizer then inserts the sentence at the beginning of the document before the start of the text...

...analysis to generate a list of ranked sentences for consideration in the summary. The summarizer counts how frequently content words appear in a document and produces a table correlating the content words with their corresponding frequency...

...inclusion of the sentence have been satisfied. The summarizer then inserts the sentence at the beginning of the document before the start of the text. >

Claims:

...based summary of a document's writings; and inserting the sentence-based summary at a beginning of the document.

...

...frequently words appear in the document, a computer-implemented method comprising: evaluating words in the document to identify ordered sets of words that appear repeatedly in a same order; ranking individual sentences in the document by treating the ordered sets of words as if they were single words; generating the summary based at least in part on the sentence rankings; inserting the summary into a file comprising the document; and saving the file to non-volatile data storage.

11/5,K/3 (Item 3 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2008 The Thomson Corporation. All rts. reserv.

0010977638 - Drawing available
WPI ACC NO: 2001-601873/200168
Related WPI Acc No: 1999-404693; 2002-656541; 2006-687590
XRPX Acc No: N2001-449060

Computer-based document summarization method for word processor, involves performing cue-phrase analysis by comparing words and phrases in specified sentences with pre-compiled list

Patent Assignee: COKUS S J (COKU-I); DOLAN W B (DOLA-I); FEIN R A (FEIN-I); FRIES E J (FRIE-I); MESSERLY J (MESS-I); MICROSOFT CORP (MICT); THORPE C A (THOR-I)

Inventor: COKUS S J; DOLAN W B; FEIN R A; FRIES E J; MESSERLY J; THORPE C A
Patent Family (2 patents, 1 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update
US 20010021938	A1	20010913	US 1996622864	A	19960329	200168 B
			US 1999289085	A	19990408	
US 6349316	B2	20020219	US 1999289085	A	19990408	200221 E

Priority Applications (no., kind, date): US 1996622864 A 19960329; US 1999289085 A 19990408

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 20010021938	A1	EN	12	3	Continuation of application US
1996622864					

Continuation of patent US 5924108

Alerting Abstract US A1

NOVELTY - The individual sentences are scored with the corresponding rankings according to respective frequency of content words. A cue-phrase analysis is performed by comparing words and phrases in sentences with the pre-compiled list. A summary is created based on comparison result.

DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

- 1.Word processing application;
- 2.Electronic mail application;
- 3.Internet web browser application;
- 4.Programmed computer for document summarizing;
- 5.Document file

USE - For summarizing documents for word processing, electronic mail and Internet web browser applications.

ADVANTAGE - Enables the author to automatically create summaries using the statistical and cue-pharse approach as it is designed from the author's standpoint and to place the created summaries conveniently at the top of the document. Improves the quality of the final summary.

DESCRIPTION OF DRAWINGS - The figure shows the computer loaded with word processing program for performing document summarizer function.

Title Terms/Index Terms/Additional words: COMPUTER; BASED; DOCUMENT; METHOD ; WORD; PROCESSOR; PERFORMANCE; CUE; PHRASE; ANALYSE; COMPARE; SPECIFIED; SENTENCE; PRE; COMPILE; LIST

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0017/27 A I R 20060101

G06F-0017/30 A I R 20060101

G06F-0017/27 C I R 20060101

G06F-0017/30 C I R 20060101

US Classification, Issued: 707531, 707500, 707531

File Segment: EPI;

DWPI Class: T01

Manual Codes (EPI/S-X): T01-E01C; T01-H07C1; T01-H07C3C; T01-H07C5E;

T01-J11A1; T01-J12B

Original Publication Data by Authority

Original Abstracts:

...summarizer performs a statistical analysis to generate a list of ranked sentences for consideration in the summary. The summarizer counts how frequently content words appear in a document and produces a table correlating the content words with their corresponding frequency counts. Phrase compression techniques...

...inclusion of the sentence have been satisfied. The summarizer then inserts the sentence at the beginning of the document before the start of the text...

...performs a statistical analysis to generate a list of ranked sentences for consideration in the summary. The summarizer counts how frequently content words appear in a document and produces a table correlating the content words with their corresponding frequency counts. Phrase compression techniques are used to produce more accurate counts...

...inclusion of the sentence have been satisfied. The summarizer then inserts the sentence at the beginning of the document before the start of the text.

Claims:

...A computer-implemented method for summarizing documents, comprising the following steps: counting how frequently content words appear in a document; scoring individual sentences according to their respective content words, wherein sentences which contain more content words that appear more frequently in the document are...

...appear in a document to produce frequency counts for corresponding content words; (b) scoring individual sentences according to the content words contained in the sentences; (c) identifying a sentence with the highest score; (d) adjusting the frequency counts of the content

words that appear in the highest scoring sentence to remove an influence of the highest scoring sentence ; and (e) re-scoring the sentences based on the adjusted frequency counts.

11/5,K/4 (Item 4 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2008 The Thomson Corporation. All rts. reserv.

0009869497 - Drawing available
WPI ACC NO: 2000-165378/200015
XRPX ACC No: N2000-123876

Computer aided text translation apparatus used in Japanese to English translation - has Japanese and English text search units which search and perform translation of input search sentence, based on comparison of both languages text stored in bilingual sentence storing unit

Patent Assignee: FUJI XEROX CO LTD (XERF)
Inventor: MASUICHI H; TATENO M; TATENO S; UMEKI H; UMEMOTO H
Patent Family (3 patents, 2 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	
JP 2000020524	A	20000121	JP 1998202788	A	19980702	200015	B
JP 3114703	B2	20001204	JP 1998202788	A	19980702	200065	E
US 6321189	B1	20011120	US 1999343543	A	19990630	200174	E

Priority Applications (no., kind, date): JP 1998202788 A 19980702

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
JP 2000020524	A	JA	18	8	
JP 3114703	B2	JA	15		Previously issued patent JP 2000020524

Alerting Abstract JP A

NOVELTY - A search sentence input unit (2) receives text to be translated from Japanese language to English. Based on comparison of both the languages, Japanese and English text search units (3,4) search and perform the translation of the search sentence. DETAILED DESCRIPTION - A bilingual sentence storing unit (1) stores a sentence written in Japanese and its translations are written in English correspondingly.

USE - Used in Japanese to English translation.

ADVANTAGE - Suitable bilingual sentence search operation is performed, without depending on variant in expression of Japanese language. Unnecessitates need of dictionary to be produced beforehand. Since correspondence relationship between two languages are dynamically acquired from extensive bilingual sentence pair information, effect of which bilingual sentence search is performed depending on the input search question. DESCRIPTION OF DRAWING(S) - The figure shows the components of typical computer aided translation apparatus. (1) Bilingual sentence storing unit; (2) Search sentence input unit; (3,4) Japanese and English text search units.

Title Terms/Index Terms/Additional words: COMPUTER; AID; TEXT; TRANSLATION; APPARATUS; JAPAN; ENGLISH; SEARCH; UNIT; PERFORMANCE; INPUT; SENTENCE; BASED; COMPARE; LANGUAGE; STORAGE; BILINGUAL

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0017/27	A	I	R	20060101
G06F-0017/28	A	I	R	20060101
G06F-0017/30	A	I	R	20060101
G06F-0017/27	C	I	R	20060101
G06F-0017/28	C	I	R	20060101
G06F-0017/30	C	I	R	20060101

US Classification, Issued: 7047, 7048, 7075, 707536

File Segment: EPI;

DWPI Class: T01

Manual Codes (EPI/S-X): T01-E01C; T01-J05B3; T01-J14

Original Publication Data by Authority

Claims:

...similar to the query from among a set of the first language sentences stored in the pair **data** storing means; and second retrieving means for retrieving second language sentences similar to second language...

...retrieving means from among a set of the second language sentences stored in the pair **data** storing means; wherein the second retrieving means determines and extracts important words from each of...

...importance, wherein, for the set A of the second language sentences stored in the pair **data** storing means, a set B of the second language sentences having the same meaning and paired with the respective first language sentences retrieved by the first retrieving means and a set C of all words appearing in the set B, a first value that is the number of sentences included in the set B, a second value that is the number of sentences in the set B for each important word candidate containing the important word candidate, supposing...

11/5,K/5 (Item 5 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2008 The Thomson Corporation. All rts. reserv.

0009464344 - Drawing available
WPI ACC NO: 1999-404693/199934
Related WPI Acc No: 2001-601873; 2002-656541; 2006-687590
XRPX Acc No: N1999-301661

Document summarizer method for word processor

Patent Assignee: MICROSOFT CORP (MICT)
Inventor: CHRISTOPHER; COKUS S J; DOLAN W B; EDWARD; FEIN R A; FRIES E J;
JOHN M; MESSERLY J; RONALD; SHAWN; THORPE C A; WILLIAM

Patent Family (2 patents, 2 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	
US 5924108	A	19990713	US 1996622864	A	19960329	199934	B
JP 11259457	A	19990924	JP 199840650	A	19980223	199951	NCE

Priority Applications (no., kind, date): US 1996622864 A 19960329; JP 199840650 A 19980223

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 5924108	A	EN	11	3	
JP 11259457	A	JA	49		

Alerting Abstract US A

NOVELTY - The content words are compared with precompiled list of words, which sets range conditions. A summary (72) is created which contains higher ranked sentences and condition satisfying condition.

DESCRIPTION - The frequency of the content words in a document (70) are counted and the sentence which contain more number of content words are ranked higher than those sentences which contain fewer high frequency content words. An INDEPENDENT CLAIM is also included for the computer for summarizing documents.

USE - For summarizing documents in word processors, electronic mail, internet web browser e.g. internet explorer from microsoft corporation.

ADVANTAGE - Creates summaries using combined statistical and cue-phrase approach thus improving the quality of the summary. Enables author to place the summary at the top of the document, facilitating the author to revise the summary as per his wish.

DESCRIPTION OF DRAWINGS - The figure shows the documents with summaries created.

70 Document

72 Summary

Title Terms/Index Terms/Additional words: DOCUMENT; METHOD; WORD; PROCESSOR

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0017/21 A I F R 20060101

G06F-0017/27 A I R 20060101
G06F-0017/30 A I L R 20060101
G06F-0017/30 A I R 20060101
G06F-0017/21 C I F R 20060101
G06F-0017/27 C I R 20060101
G06F-0017/30 C I L R 20060101
G06F-0017/30 C I R 20060101
US Classification, Issued: 707531

File Segment: EPI;
DWPI Class: T01
Manual Codes (EPI/S-X): T01-J11A

Original Publication Data by Authority

Original Abstracts:

...Summarizer performs a statistical analysis to generate a list of ranked sentences for consideration in the summary. The summarizer counts how frequently content words appear in a document and produces a table correlating the content words with their corresponding frequency counts. Phrase compression techniques...

...inclusion of the sentence have been satisfied. The summarizer then inserts the sentence at the beginning of the document before the start of the text.

21/5,K/1 (Item 1 from file: 347)
DIALOG(R)File 347:JAPIO
(c) 2008 JPO & JAPIO. All rts. reserv.

07055140 **Image available**
INFORMATION PROCESSOR AND PRINTER

PUB. NO.: 2001-282775 [JP 2001282775 A]
PUBLISHED: October 12, 2001 (20011012)
INVENTOR(s): TADA TOMOYUKI
APPLICANT(s): OMRON CORP
APPL. NO.: 2000-090054 [JP 200090054]
FILED: March 29, 2000 (20000329)
INTL CLASS: G06F-017/21; B41J-021/00; G06F-003/12

ABSTRACT

PROBLEM TO BE SOLVED: To calculate the **quantity** of **sentences** corresponding to personal information, and to select and lay out an image in the size suited to a margin thereof.

SOLUTION: On the basis of data registered in a personal attribute data base 32 and a suited condition data base 33, print **information** is generated by a print information synthesizing processing part 31 and outputted to a print layout processing part 34. The print layout processing part 34 reads sentences matched to the print information from a print **sentence** data base 34, calculates the **quantity** of print **sentences**, calculates the **quantity** of margin on the basis of the calculated result of **sentence quantity**, retrieves the image ID of the image in the optimal size from a print image data base 37 and selects a correspondent image out of an image **data group** 38 on the basis of that image ID. While referring to a print layout data base 35, the print layout is determined and outputted to a print data output part 39 and data are made to be printed.

COPYRIGHT: (C)2001,JPO

...PUBLISHED: 20011012)

ABSTRACT

PROBLEM TO BE SOLVED: To calculate the **quantity** of **sentences** corresponding to personal information, and to select and lay out an image in the size...

... of data registered in a personal attribute data base 32 and a suited condition data base 33, print **information** is generated by a print information synthesizing processing part 31 and outputted to a print...
... print layout processing part 34 reads sentences matched to the print information from a print **sentence** data base 34, calculates the **quantity** of print **sentences**, calculates the **quantity** of margin on the basis of the calculated result of **sentence quantity**, retrieves the image ID of the image in the optimal size from a print image data base 37 and selects a correspondent image out of an image **data group** 38 on the basis of that image ID. While referring to a print layout data...

21/5,K/2 (Item 2 from file: 347)
DIALOG(R)File 347:JAPIO
(c) 2008 JPO & JAPIO. All rts. reserv.

04408262 **Image available**
DOCUMENT STRUCTURE DATA BASE CONSTRUCTION PROCESSING SYSTEM

PUB. NO.: 06-052162 [JP 6052162 A]
PUBLISHED: February 25, 1994 (19940225)
INVENTOR(s): HOSHIAI TADASHI
APPLICANT(s): FUJITSU LTD [000522] (A Japanese Company or Corporation), JP (Japan)
APPL. NO.: 03-241403 [JP 91241403]
FILED: September 20, 1991 (19910920)
INTL CLASS: [5] G06F-015/20; G06F-015/20

JAPIO CLASS: 45.4 (INFORMATION PROCESSING -- Computer Applications)
JOURNAL: Section: P, Section No. 1746, Vol. 18, No. 287, Pg. 128, May
31, 1994 (19940531)

ABSTRACT

PURPOSE: To automatically construct the data base of a document structure network on a document structure data base construction processing system for constructing the data base of the document structure network describing the paragraph constitution of a document.

CONSTITUTION: A paragraph sentence separation part 11 segmenting text data on the input document in the unit of a paragraph sentence, a paragraph class feature quantity management part 14 managing a feature quantity which a paragraph class classifying the paragraph sentence has, a document structure specification part 15 specifying the paragraph class to which the respective paragraph sentences segmented by the paragraph sentence separation part 11 belong by referring to management data of the paragraph class feature quantity management part 14, and specifying the document structure network of the input document by specifying connection between the paragraph classes and a document structure management part 17 managing the document structure network specified by the document structure specification part 15 for the respective document businesses are provided.

...PUBLISHED: 19940225)

ABSTRACT

PURPOSE: To automatically construct the data base of a document structure network on a document structure data base construction processing system for constructing the data base of the document structure network describing the paragraph constitution of a document...

... part 11 segmenting text data on the input document in the unit of a paragraph sentence, a paragraph class feature quantity management part 14 managing a feature quantity which a paragraph class classifying the paragraph sentence...

21/5,K/3 (Item 1 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2008 The Thomson Corporation. All rts. reserv.

0015728921 - Drawing available

WPI ACC NO: 2006-290811/200630

XRPX Acc No: N2006-247728

Semantic annotation providing method for use in data processing system, involves dividing data set of sentences into set of corpuses, and learning structure of each sentence of corpus using set of trainers

Patent Assignee: INT BUSINESS MACHINES CORP (IBMC)

Inventor: GAO Y; PICHENY M A; SARIKAYA R

Patent Family (1 patents, 1 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update
US 20060074634	A1	20060406	US 2004959523	A	20041006	200630 B

Priority Applications (no., kind, date): US 2004959523 A 20041006

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 20060074634	A1	EN	14	8	

Alerting Abstract US A1

NOVELTY - The method involves dividing a data set of sentences into a set of corpuses, where each of the corpuses comprises an equal number of sentences. A structure of each sentence of the corpus is learned using a set of trainers, and a model is formed based on the structure. A new sentence is annotated using the model in a set of engines, and the model is trained using a parse tree that is annotated by a human annotator

DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

- 1.a framework for fast semi-automatic semantic annotation, comprising an annotation tool
- 2.a data processing system for fast semi-automatic semantic annotation,

comprising a dividing unit

3.a computer program product in a computer readable medium, comprising a set of instructions for dividing a data set of sentences into a set of corpuses .

USE - Used for providing a semantic annotation in a data processing system (claimed), to train an initial parser.

ADVANTAGE - The method increases amount of training data provided for each round of annotation, so that the parser learns more and makes fewer mistakes in annotation each time, and hence minimizing time and cost of human annotation required for inspecting and correcting annotated sentences, thereby reducing efforts required for human annotation.

DESCRIPTION OF DRAWINGS - The drawing shows a block representation of a semantically annotated sentence.

522 Pron-sub

524 Subject

530 Intend

540 verb

550 City

Title Terms/Index Terms/Additional Words: METHOD; DATA; PROCESS; SYSTEM; DIVIDE; SET; SENTENCE; LEARNING; STRUCTURE; CORPUS; TRAINING

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0017/27 A I F B 20060101

US Classification, Issued: 7049

File Segment: EPI;

DWPI Class: T01

Manual Codes (EPI/S-X): T01-J16C3; T01-S03

Semantic annotation providing method for use in data processing system, involves dividing data set of sentences into set of corpuses, and learning structure of each sentence of corpus using set of trainers

Alerting Abstract ...NOVELTY - The method involves dividing a data set of sentences into a set of corpuses, where each of the corpuses comprises an equal number of sentences . A structure of each sentence of the corpus is learned using a set of trainers...

...comprising a dividing unit a computer program product in a computer readable medium, comprising a set of instructions for dividing a data set of sentences into a set of corpuses .

...

...Used for providing a semantic annotation in a data processing system (claimed), to train an initial parser.

Original Publication Data by Authority

Claims:

...is: 1. A method in a data processing system for fast semi-automatic semantic annotation, the method comprising: dividing a data set of sentences into a plurality of corpuses, wherein each of the plurality of corpuses includes an equal number of sentences ; learning a structure of each sentence of a first corpus using a plurality of trainers; forming a model based on the structure; and using the model...

Basic Derwent Week: 200630

21/5,K/4 (Item 2 from file: 350)

DIALOG(R)File 350:Derwent WPIX

(c) 2008 The Thomson Corporation. All rts. reserv.

0015728917 - Drawing available

WPI ACC NO: 2006-290807/200630

XRPX Acc No: N2006-247724

Statistical classifier constructing method, involves constructing binary valued feature vectors for sentences in training data, and calculating initial word/class probability parameter values based on training data feature vectors

Patent Assignee: MICROSOFT CORP (MICT)

Inventor: ACERO A; CHELBA C

Patent Family (1 patents, 1 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update
US 20060074630	A1	20060406	US 2004941399	A	20040915	200630 B

Priority Applications (no., kind, date): US 2004941399 A 20040915

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 20060074630	A1	EN	18	8	

Alerting Abstract US A1

NOVELTY - The method involves receiving labeled training data (308) with text sentences labeled by a class, and calculating an initial class probability parameter values for each class based on the training data sentences. A set of binary valued feature vectors (318) are constructed for sentences in the training data, and an initial word/class probability parameter values are calculated based on the training data feature vectors.

DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

- 1.a computer readable medium including instructions readable by a computer for constructing a statistical classifier
- 2.a speech utterance classification system comprising a speech utterance classification engine.

USE - Used for constructing a statistical classifier on a natural language interface.

ADVANTAGE - The method effectively improves the model performance of a statistical classifier with a faster convergence speed.

DESCRIPTION OF DRAWINGS - The drawing shows a block diagram of a system for constructing a statistical classifier.

- 302 Classifier construction module
- 308 Training data
- 310 Pre-processing module
- 318 Binary valued feature vector
- 322 Initialization module

Title Terms/Index Terms/Additional words: STATISTICAL; CLASSIFY; CONSTRUCTION; METHOD; BINARY; VALUE; FEATURE; VECTOR; SENTENCE; TRAINING; DATA; CALCULATE; INITIAL; WORD; CLASS; PROBABILITY; PARAMETER; BASED

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0017/27 A I F B 20060101

US Classification, Issued: 7049

File Segment: EPI;

DWPI Class: T01

Manual Codes (EPI/S-X): T01-J05C; T01-J18; T01-S03

...and calculating an initial class probability parameter values for each class based on the training data sentences. A set of binary valued feature vectors (318) are constructed for sentences in the training data, and...

Original Publication Data by Authority

Claims:

...calculating an initial class probability parameter thetaya values for each class y based on the number of training data sentences having the corresponding class label;constructing a set of binary valued feature

vectors for sentences in the training data , each set of feature vectors corresponding to a class label, each feature vector corresponding to a sentence, each feature corresponding to a word k;calculating initial word/class probability parameter thetak values based on training data feature...
Basic Derwent Week: 200630

21/5,K/5 (Item 3 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2008 The Thomson Corporation. All rts. reserv.

0015552291 - Drawing available
WPI ACC NO: 2006-116446/200612
XRPX Acc No: N2006-100822

Word boundary probability estimating device for use in e.g. spelling checking, has estimator estimating probability of boundary existing in set of characters by referring to calculated probability between another set of characters

Patent Assignee: IBM CORP (IBMC); INT BUSINESS MACHINES CORP (IBMC)
Inventor: MORI S; TAKUMA D

Patent Family (2 patents, 2 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	
US 20060015326	A1	20060119	US 2005180153	A	20050713	200612	B
JP 2006031295	A	20060202	JP 2004207864	A	20040714	200612	E

Priority Applications (no., kind, date): JP 2004207864 A 20040714

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 20060015326	A1	EN	19	12	
JP 2006031295	A	JA	19		

Alerting Abstract US A1

NOVELTY - The device has a calculator for calculating a probability of a word boundary existing between a set of characters that constitute a character string stored in a corpus by invoking information. An estimator estimates the probability of the boundary existing in another set of characters that constitute another string stored in another corpus by referring to the calculated probability between the latter set of characters.

DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

- 1.an unknown word model building method;
- 2.a word boundary probability estimating method;
- 3.stored software.

USE - For estimating a probability of a word boundary, in kana-kanji conversion, spelling checking, optical character recognition and speech recognition technique.

ADVANTAGE - The configuration of the device improves the accuracy of recognition in natural language processing.

DESCRIPTION OF DRAWINGS - The drawing shows a kana-kanji converting device.

- 22 Language decoding section
- 30 Base form pool
- 300 vocabulary dictionary
- 302 Character dictionary
- 320, 322 Corpuses

Title Terms/Index Terms/Additional words: WORD; BOUNDARY; PROBABILITY; ESTIMATE; DEVICE; SPELLING; CHECK; EXIST; SET; CHARACTER; REFER; CALCULATE

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0017/27	A	I	F	B	20060101
G10L-0015/18	A	I	L	B	20060101

G10L-0015/00 C I L B 20060101
US Classification, Issued: 7049

File Segment: EPI;
DWPI Class: T01
Manual Codes (EPI/S-X): T01-C08A; T01-J05B1; T01-J14; T01-J16C2; T01-S03
Original Publication Data by Authority

Original Abstracts:

...a relatively large corpus, are given as a training corpus that is storage containing vast quantities of sample sentences. Vocabulary including contextual information is expanded from words occurring in first corpus of relatively small size to words occurring...

Claims:

...containing the first character string comprising the first plurality of characters or setting up the preliminary information as to whether the word boundary exists, and means for estimating the probability that the word boundary will exist in a second plurality of characters constituting

Basic Derwent Week: 200612

21/5,K/6 (Item 4 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2008 The Thomson Corporation. All rts. reserv.

0015067516 - Drawing available
WPI ACC NO: 2005-416751/200542
XRPX Acc No: N2005-338098

Sentences generating method for use by e.g. writer, involves generating list of words for each word present in input source from attached repositories in particular language, and combining all generated lists to generate sentences

Patent Assignee: BEHBEHANI H (BEHB-I)

Inventor: BEHBEHANI H

Patent Family (1 patents, 1 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update
US 20050120002	A1	20050602	US 2003507518	P	20031002	200542 B
			US 2004939353	A	20040914	

Priority Applications (no., kind, date): US 2003507518 P 20031002; US 2004939353 A 20040914

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 20050120002	A1	EN	11	3	Related to Provisional US 2003507518

Alerting Abstract US A1

NOVELTY - The method involves analyzing a source text, and extracting words from a source. A list of words is generated for each word present in an input source from attached repositories in a particular language based on desired retrieval mechanism such as predefine lists, aliases and synonyms. The list is displayed, and a set of desired words is selected from the list. All the generated lists are combined to generate sentences.

DESCRIPTION - An INDEPENDENT CLAIM is also included for a service for generating sentences.

USE - Used for generating sentences by publishing organization, writer, author, lecturer, teacher and institution to provide textbook and courseware according to students interests and details, medical research doctor for combining medicines, diseases and symptoms, and for UPSTO employee, USPTO customer, artistic research, creative activity and brain storming.

ADVANTAGE - The method facilitates searching of custom repositories such as documents and databases, in an easy manner. The method helps users who want to search some repositories which are not in their respective languages. The method is useful not only for corporate entities but also for individuals.

DESCRIPTION OF DRAWINGS - The drawing shows an inner working of a process of generating sentences.

Title Terms/Index Terms/Additional words: SENTENCE; GENERATE; METHOD;
WRITING; LIST; WORD; PRESENT; INPUT; SOURCE; ATTACH; LANGUAGE;
COMBINATION

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0017/28 A I R 20060101

G06F-0017/28 C I R 20060101

US Classification, Issued: 7073

File Segment: EPI;

DWPI Class: T01

Manual Codes (EPI/S-X): T01-J11A

Original Publication Data by Authority

Original Abstracts:

...process of text generation/creation is automated. The text to be processed is used as **seed** for the **text generation** process. The **text** to be processed can be in any language and can be passed to text generation

Claims:

...Text means data in specific format. It may be single or multiple sentences, characters, words, **numbers**, formulae and expressions. .. etc **sentences** and text are used alternatively. Alias means a unique name for accessing multiple lists. An alias can be created by **combining** multiple predefined lists thereby **and** all the entries are all the lists are accessed attached with a alias.output device...

Basic Derwent week: 200542

21/5,K/7 (Item 5 from file: 350)

DIALOG(R)File 350:Derwent WPIX

(c) 2008 The Thomson Corporation. All rts. reserv.

0013881346 - Drawing available

WPI ACC NO: 2004-060250/ 200406

SRPX Acc No: N2004-048721

Computerized medical information management system for hospital, uses set of universal templates for entering medical data and intelligent data fields that adapt to user input automatically

Patent Assignee: BURSTEIN A (BURS-I); BURSTEIN B (BURS-I)

Inventor: BURSTEIN A; BURSTEIN B

Patent Family (1 patents, 1 countries)

Patent Application

Number	Kind	Date	Number	Kind	Date	Update
US 20030220819	A1	20031127	US 2002151155	A	20020521	200406 B

Priority Applications (no., kind, date): US 2002151155 A 20020521

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 20030220819	A1	EN	23	14	

Alerting Abstract US A1

NOVELTY - The management software uses a minimum set of intelligent templates that display medical conditions in graphical user interface (GUI) and intelligent data fields that adapt to user input automatically. The system includes modules for tracking patients, indicating occupancy of bed, and to output billing statements, insurance/management reports, and additional administrative documents in grammatically accurate, understandable phrases.

DESCRIPTION - An INDEPENDENT CLAIM is also included for computer program product comprising readable medium storing instructions for managing medical information.

USE - In hospital for medical management using intranet and Internet.

ADVANTAGE - Eliminates transcription and simplifies the task of recording symptoms diagnosis and patient history.

DESCRIPTION OF DRAWINGS - The figure shows a close-up, partial screen shot of the on-screen data entry form.

Title Terms/Index Terms/Additional words: COMPUTER; MEDICAL; INFORMATION; MANAGEMENT; SYSTEM; HOSPITAL; SET; UNIVERSAL; TEMPLATE; ENTER; DATA; INTELLIGENCE; FIELD; ADAPT; USER; INPUT; AUTOMATIC

Class Codes

International Classification (Main): G06F-017/60
US Classification, Issued: 7053

File Segment: EPI;

DWPI Class: S05; T01

Manual Codes (EPI/S-X): S05-G02G1; T01-J06A1; T01-J12B; T01-N02A2; T01-S03

Computerized medical information management system for hospital, uses set of universal templates for entering medical data and intelligent data fields that adapt to user...

Original Publication Data by Authority

Claims:

...for executing program code under the direction of the processor, a storage device for storing data and program code and a bus connecting the processor and the storage device; e) a...

...means for creating, from data entered by the user, reports in natural language consisting of grammatical, readily understood sentences and phrases; 3) intelligent data fields that adapt according to data previously entered by the...

21/5,K/8 (Item 6 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2008 The Thomson Corporation. All rts. reserv.

0013732622 - Drawing available
WPI ACC NO: 2003-830693/ 200377
Related WPI Acc No: 2006-619768
XRPX Acc No: N2003-663756

Webpage reading method for sight-impaired users, involves reading webpage from initial reading position according to user- configurable settings

Patent Assignee: INT BUSINESS MACHINES CORP (IBMC)

Inventor: CRAGUN B J

Patent Family (2 patents, 1 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	
US 20030172353	A1	20030911	US 200293159	A	20020307	200377	B
US 7058887	B2	20060606	US 200293159	A	20020307	200638	E

Priority Applications (no., kind, date): US 200293159 A 20020307

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 20030172353	A1	EN	21	10	

Alerting Abstract US A1

NOVELTY - The method involves determining a set of user-configurable settings for reading the webpage. An initial reading position on the webpage is determined based on the user-configurable settings. The webpage is then read from the initial reading position according to the user-configurable settings.

DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

- 1.computer readable medium containing program for reading webpage; and
- 2.computer program product for reading webpage.

USE - For programmatically reading webpage using personal digital assistant (PDA), wireless device for sight-impaired users.

ADVANTAGE - By selecting initial display position in a document, optimum use of devices with limited display area and communication bandwidth is enabled.

DESCRIPTION OF DRAWINGS - The figure the flowchart illustrating the operation of webpage reader program.

Title Terms/Index Terms/Additional words: READ; METHOD; SIGHT; IMPAIR; USER ; INITIAL; POSITION; ACCORD; CONFIGURATION; SET

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0015/00 A I F B 20060101

G06F-0015/00 A I R 20060101

G06F-0015/00 C I L B 20060101

G06F-0015/00 C I R 20060101

US Classification, Issued: 715517, 715517, 715523, 704270.1, 709201

File Segment: EPI;

DWPI Class: S05; T01

Manual Codes (EPI/S-X): S05-K; T01-J11C; T01-N03A1; T01-S03

Original Publication Data by Authority

Original Abstracts:

A method and apparatus for reading a web page according to a set of user-configurable settings. In one embodiment, a set of user-configurable settings configured for reading the web page is determined. An initial reading position on the web page is determined as specified by the user-configurable settings. The web page is then read...

...A method and apparatus for reading a web page according to a set of user-configurable settings. In one embodiment, a set of user-configurable settings configured for reading the web page is determined. An initial reading position on the web page is determined as specified by the user-configurable settings. The web page is then read from the initial reading position according to the set of user-configurable settings.

Claims:

what is claimed is: 1. A method of reading a web page according to a set of user-configurable settings, comprising: determining the set of user-configurable settings; determining an initial reading position on the web page as specified by the set of user-configurable settings; and reading, by a reading program, the web page from the initial reading position according to the set of user-configurable settings.

...

...what is claimed is: 1. A computer-implemented method of reading a web page according to a predefined set of user-configurable settings, comprising: upon retrieving the web page, selecting a setting from the set of user-configurable settings on the basis of an attribute of the web page, wherein the attribute is at least one of content of the web page and a URL of the web page; determining an initial reading position on the web page as specified by the selected setting; and reading, by a reading program, the web page from the initial reading position according to the set of user-configurable settings, wherein the user-configurable settings are at least one of: a URL setting configured to identify the web page on the basis of the URL; a link page setting configured to identify the web page as a link page dependent on a quantification...

...to identify the web page as an overview page dependent on a quantification of a number of sentences of readable text in the web page. Basic Derwent Week: 200377

21/5,K/9 (Item 7 from file: 350)

DIALOG(R)File 350:Derwent WPIX

(c) 2008 The Thomson Corporation. All rts. reserv.

0013615310 - Drawing available

WPI ACC NO: 2003-710624/ 200367

XRPX Acc No: N2003-568202

Redundant information removal method for digital document, involves organizing document text into sentences and paragraphs, and comparing organized document with other documents to identify redundancies

Patent Assignee: BOREK S E (BORE-I); BOURBAKIS N G (BOUR-I); US SEC OF AIR FORCE (USAF)

Inventor: BOREK S E; BOURBAKIS N G

Patent Family (2 patents, 1 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	
US 20030145279	A1	20030731	US 2002351636	P	20020125	200367	B
			US 2002314189	A	20021205		
US 7017113	B2	20060321	US 2002314189	A	20021205	200621	E

Priority Applications (no., kind, date): US 2002351636 P 20020125; US 2002314189 A 20021205

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 20030145279	A1	EN	12	6	Related to Provisional US 2002351636

Alerting Abstract US A1

NOVELTY - The text of original digital document retrieved from the database (100) is organized into sentences and paragraphs. The organized document is analyzed and compared with other documents to identify the redundancies present in the documents, using information redundancy removal (IRR) software (140).

DESCRIPTION - An INDEPENDENT CLAIM is also included for apparatus for removing redundant information from digital document.

USE - For removing redundant information such as paragraph of text or images from original document such as web pages, for reconstruction of new document related to government organization.

ADVANTAGE - Removes redundant information from retrieved documents using simple technique.

DESCRIPTION OF DRAWINGS - The figure shows the block diagram of the redundant information removal process.

100 database

120 search engine

140 IRR software

180 new document

Title Terms/Index Terms/Additional words: REDUNDANT; INFORMATION; REMOVE; METHOD; DIGITAL; DOCUMENT; ORGANISE; TEXT; SENTENCE; PARAGRAPHS; COMPARE; IDENTIFY

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0017/00 A I F B 20060101

G06F-0017/27 A I R 20060101

G06F-0017/30 A I R 20060101

G06F-0017/00 C I L B 20060101

G06F-0017/27 C I R 20060101

G06F-0017/30 C I R 20060101

US Classification, Issued: 715511, 715530, 715534

File Segment: EPI;

DWPI Class: T01

Manual Codes (EPI/S-X): T01-J05B4P; T01-J11A1; T01-J11C; T01-N03A2

Original Publication Data by Authority

Original Abstracts:

Method and apparatus for reconstructing new documents from a group of old ones by removing the existing redundant information. Redundant information (images, text paragraphs) from retrieved multimedia documents is removed. Each document consists of two main parts stored in different databases. The first part of a document represents text paragraphs, the...

...paragraphs, by keeping pointers useful for a future reconstruction of

the original documents. The remaining text paragraphs and the set of points are used to compose the first version of a new document. The invention also examines all the images related with the set of original documents and removes the same or similar images while keeping pointers that could assist a future reconstruction of the original...

...Method and apparatus for reconstructing new documents from a group of old ones by removing the existing redundant information. Redundant information (images, text paragraphs) from retrieved multimedia documents is removed. Each document consists of two main parts stored in different databases. The first part of a document represents text paragraphs, the second part consists of the images and drawings related with the...

...paragraphs, by keeping pointers useful for a future reconstruction of the original documents. The remaining text paragraphs and the set of points are used to compose the first version of a new document. The invention also examines all the images related with the set of original documents and removes the same or similar images while keeping pointers that could assist a future reconstruction of the original documents. The invention merges text...

Claims:

...of a paragraph in characters; character histograms; number of words in each sentence; word histograms; starting word of each sentence; and ending word of a paragraph; determining whether similar said statistical...

...THEN deciding paragraphs are similar, removing redundant paragraph, and proceeding to said step of comparing said sentences and paragraphs with other documents OTHERWISE, postponing removal of paragraph; analyzing corresponding image and data...

Basic Derwent Week: 200367

21/5,K/10 (Item 8 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2008 The Thomson Corporation. All rts. reserv.

0009670004 - Drawing available
WPI ACC NO: 1999-623996/ 199954
XRPX Acc No: N1999-460741

Text structure analysis apparatus for documentation apparatus - has tree structure determining unit to determine tree structure, based on degree of importance calculated between sentences

Patent Assignee: SHARP KK (SHAF)

Inventor: OKUNISHI T; YAMAJI T; YOSHIMI T

Patent Family (3 patents, 2 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	
JP 11272664	A	19991008	JP 199870288	A	19980319	199954	B
US 6374209	B1	20020416	US 1999271569	A	19990318	200232	E
JP 3429184	B2	20030722	JP 199870288	A	19980319	200350	E

Priority Applications (no., kind, date): JP 199870288 A 19980319

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
JP 11272664	A	JA	17	6	
JP 3429184	B2	JA	17		Previously issued patent JP 11272664

Alerting Abstract JP A

NOVELTY - The input text is divided into sentence and stored in memory (8). Relation degree calculator (4) calculates relation between main concept and the sentence that are stored in memory. Based on their relation importance degree, calculator (5) calculates the degree of importance of the sentence, based on which tree structure of input text is determined.

DETAILED DESCRIPTION - An output unit (7) displays the obtained tree structure of the text. Essential word recognition unit (2) recognizes essential word from each row and stores it in memory. An INDEPENDENT CLAIM is also included for analysis program recording medium.

USE - For documentation apparatus.

ADVANTAGE - As text are extracted based on their degree of importance, an accurate text structure analysis is performed. DESCRIPTION OF DRAWING(S) -

The figure shows the text structure analysis apparatus. (2) word recognition unit; (4,5) Calculators; (7) Output unit; (8) Memory.

Title Terms/Index Terms/Additional words: TEXT; STRUCTURE; ANALYSE; APPARATUS; DOCUMENT; TREE; DETERMINE; UNIT; BASED; DEGREE; IMPORTANT; CALCULATE; SENTENCE

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0017/21	A	I	F	R	20060101
G06F-0017/27	A	I		R	20060101
G06F-0017/30	A	I	L	R	20060101
G06F-0017/21	C	I	F	R	20060817
G06F-0017/27	C	I		R	20060101
G06F-0017/30	C	I	L	R	20060101

US Classification, Issued: 7049, 707531

File Segment: EPI;

DWPI Class: T01

Manual Codes (EPI/S-X): T01-J05B; T01-J11A

Original Publication Data by Authority

Original Abstracts:

A text input section (1) divides an inputted text into sentences and attaches a number to each of the sentences, which is stored in a text data base together with the number. An important word recognizing section (2) generates a list of important words...

Basic Derwent Week: 199954

21/5,K/11 (Item 9 from file: 350)

DIALOG(R)File 350:Derwent WPIX

(c) 2008 The Thomson Corporation. All rts. reserv.

0009395704 - Drawing available

WPI ACC NO: 1999-331691/ 199928

XRPX Acc No: N1999-249326

Automatic music composing apparatus - composes music based on extracted music templates which have characteristic information which are in accord with input conditions

Patent Assignee: YAMAHA CORP (NIHG)

Inventor: AOKI E; SUGIURA T

Patent Family (3 patents, 2 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	
JP 11119774	A	19990430	JP 1997280848	A	19971014	199928	B
US 6075193	A	20000613	US 1998170495	A	19981013	200035	E
JP 3620240	B2	20050216	JP 1997280848	A	19971014	200513	E

Priority Applications (no., kind, date): JP 1997280848 A 19971014

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
JP 11119774	A	JA	17	15	
JP 3620240	B2	JA	22		Previously issued patent JP 11119774

Alerting Abstract JP A

NOVELTY - The search unit extracts the music template containing the characteristic information which are in accord with the conditions input by user, and composes music based on the extracted templates. DETAILED DESCRIPTION - An INDEPENDENT CLAIM is also included for recording medium which stores software for music composition.

USE - For composing music automatically.

ADVANTAGE - Extract of many templates is performed. Riot of music is securable. The odd sound caused by making number of phrases and number of nodulus in accord with input conditions is suppressed. DESCRIPTION OF DRAWING(S) - The figure shows block diagram of automatic music composing apparatus.

Title Terms/Index Terms/Additional words: AUTOMATIC; MUSIC; COMPOSE;
APPARATUS; BASED; EXTRACT; TEMPLATE; CHARACTERISTIC; INFORMATION; ACCORD;
INPUT; CONDITION

Class Codes

International Classification (Main): G10H-001/00

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G10H-0001/00 A I R 20060101

G10H-0001/00 C I R 20060101

US Classification, Issued: 84609, 84610, 84634, 84649, 84650

File Segment: EngPI; EPI;

DWPI Class: w04; P86

Manual Codes (EPI/S-X): w04-U

Original Publication Data by Authority

Original Abstracts:

...a musical template data base storing a plurality of musical templates each including a first **set of data** constituting a musical melody sample defined by a pattern of musical tone pitch progression in a pattern of rhythm to be performed for a musical piece and a second **set of data** indicative of musical features of said musical melody sample. The melody sample is constructed and...

...to define musical features for a musical piece to be composed in terms of the **number of sentences**, phrases and measures, and similarity symbols of each sentence. Comparing the structure and the features...

Claims:

An automatic music composing apparatus comprising: a template data base storing a **plurality** of musical templates each including a **set of data** defining a **musical piece**, said musical piece being subdivided into a plurality of musical segments, said **set of data** including **subsets of data** respectively defining musical **properties** of said musical segments; input means for inputting composition conditions including requirements on musical properties...

Basic Derwent Week: 199928

21/5,K/12 (Item 10 from file: 350)

DIALOG(R)File 350:Derwent WPIX

(c) 2008 The Thomson Corporation. All rts. reserv.

0009075567 - Drawing available

WPI ACC NO: 1998-496375/ 199843

XRFX Acc No: N1999-033815

Computer based generation method of thematic summary from document image - involves selecting set of thematic sentences based on their score which is implemented by value related to frequency of occurrence of thematic word image in document

Patent Assignee: XEROX CORP (XERO)

Inventor: BLOOMBERG D S; CHEN F R; TUKEY J W

Patent Family (2 patents, 2 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	
BR 199606005	A	19980901	BR 19966005	A	19961213	199843	B
US 5848191	A	19981208	US 1995572848	A	19951214	199905	ETAB

Priority Applications (no., kind, date): US 1995572848 A 19951214

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
BR 199606005	A	PT	1		
US 5848191	A	EN	25	12	

Alerting Abstract US A

The method involves analysing the document image to identify sentence boundaries and to identify multiple word image equivalence class.

Predetermined number of word image equivalence class is selected as thematic word images, the number being lesser than number of thematic sentence to be interacted.

Based on occurrence of thematic word images, in the sentences, each sentence is scored. A set of thematic sentences are selected based on the score. The score of the sentence is incremented by value related to frequency of occurrence of thematic word image in the document.

USE - For generating thematic summaries without performing character recognition.

ADVANTAGE - Produces readable and semantically correct thematic summary from document image.

Title Terms/Index Terms/Additional words: COMPUTER; BASED; GENERATE; METHOD
; SUMMARY; DOCUMENT; IMAGE; SELECT; SET; SENTENCE; SCORE; IMPLEMENT;
VALUE; RELATED; FREQUENCY; OCCUR; WORD

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06K-0009/00	A	I	R	20060101
G06K-0009/20	A	I	R	20060101
G06K-0009/68	A	I	R	20060101
G06K-0009/00	C	I	R	20060101
G06K-0009/20	C	I	R	20060101
G06K-0009/68	C	I	R	20060101

US Classification, Issued: 382229, 382170, 382180, 382199, 382206, 7041

File Segment: EPI;

DWPI Class: T01

Original Publication Data by Authority

Original Abstracts:

...into text blocks, and text lines. Using the median x-height of text blocks the **main body of text** is identified. Afterward, word image equivalence classes and sentence boundaries within the blocks of the **main body of text** are determined. The word image equivalence classes are used to identify thematic words. These, in turn are used to score the sentences within the **main body of text**, and the highest scoring sentences are selected for extraction.

Claims:

...first number of word image equivalence classes, the first number being less than a second **number of thematic sentences** to be extracted; d) **scoring** each sentence of the first multiplicity of sentences based upon occurrence of thematic word images in each sentence; and e) selecting the second **number of thematic sentences** from the first multiplicity of sentences based upon the score of each sentence.

Basic Derwent Week: 199843

21/5,K/13 (Item 11 from file: 350)

DIALOG(R)File 350:Derwent WPIX

(c) 2008 The Thomson Corporation. All rts. reserv.

0008479378 - Drawing available

WPI ACC NO: 1998-008313/ 199801

Related WPI Acc No: 1996-505759; 2001-233820

XRPX Acc No: N1998-006610

Logic circuit simulator for logic circuit defined by sentence - has sentence calculating device that carries out calculation on one of number of sentences one at time and output result of calculation based on one of operator

Patent Assignee: NEC CORP (NIDE)

Inventor: TAKASAKI S

Patent Family (1 patents, 1 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update
US 5689683	A	19971118	US 1990486705	A	19900228	199801 B
			US 199374725	A	19930610	
			US 1995432260	A	19950501	

Priority Applications (no., kind, date): JP 198948225 A 19890228; JP 1989131079 A 19890524; JP 1989166926 A 19890630; JP 1989318102 A 19891207

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 5689683	A	EN	25	18	Continuation of application US 1990486705
					Division of application US 199374725
					Division of patent US 5572708

Alerting Abstract US A

The system includes a model memory for memorising a number of operators which are for carrying out operations specified by the sentences. A variable memory memorises a number of initial values of the variables specified by the sentences. A sentence calculating device is connected to the model memory and the variable memory to carry out calculation on one of the **number of sentences** one at a time and output a result of the calculation based on at least one of the operators and at least two of the initial values of the variables for each calculation of the **number of sentences**.

A data memory is connected to the sentence calculating device to memorise the results of the calculations for the **number of sentences**. A substituting device is connected to the sentence calculating device and the data memory to substitute the result of calculation for a previous result that was previously calculated according to the one of the sentences.

ADVANTAGE - Capable of dealing with description of functional level.

Title Terms/Index Terms/Additional words: LOGIC; CIRCUIT; SIMULATE; DEFINE; SENTENCE; CALCULATE; DEVICE; CARRY; ONE; NUMBER; TIME; OUTPUT; RESULT; BASED; OPERATE

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0017/50 A I R 20060101

G06F-0017/50 C I R 20060101

US Classification, Issued: 395500, 364489, 364578

File Segment: EPI;

DWPI Class: T01

Manual Codes (EPI/S-X): T01-G06; T01-J15A

...has sentence calculating device that carries out calculation on one of **number of sentences** one at time and output result of calculation based on one of operator

Alerting Abstract ...the model memory and the variable memory to carry out calculation on one of the **number of sentences** one at a time and output a result of the calculation based on at least...

...at least two of the initial values of the variables for each calculation of the **number of sentences**.

...

...connected to the sentence calculating device to memorise the results of the calculations for the **number of sentences**. A substituting device is connected to the sentence calculating device and the data memory to

Original Publication Data by Authority

Original Abstracts:

...are related to the current sentence. A data or result memory memorizes previous data or **initial result values** calculated before calculation of the current sentence. The result of calculation is substituted for those of the previous data or the **initial result values** which are related to the current sentences. Preferably, a flag memory is used to...

Basic Derwent week: 199801

21/5,K/14 (Item 12 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2008 The Thomson Corporation. All rts. reserv.

0008417803 - Drawing available
WPI ACC NO: 1997-535337/ 199749
XRPX ACC No: N1997-445754

Anticipated meaning natural language interface system for computer application - involves anticipating general meaning of each of number of likely user input sentences and storing in computer number of general meaning nodes, one for each anticipated user input general meaning

Patent Assignee: CONRAD D (CONR-I); COSBY C (COSB-I)

Inventor: CONRAD D; COSBY C

Patent Family (1 patents, 1 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update
US 5682539	A	19971028	US 1994315240	A	19940929	199749 B

Priority Applications (no., kind, date): US 1994315240 A 19940929

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 5682539	A	EN	32	28	

Alerting Abstract US A

The method involves anticipating the general meaning of each of a number of likely user input sentences and storing in the computer a number of general meaning nodes, one for each anticipated user input general meaning. Each node is associated with a function, at least one typical anticipated user-input sentence which conveys the general meaning of the node is entered. A pattern is generated from the words of the typical sentence, and the typical sentence pattern is stored in the computer. A user an input sentence is received and a pattern is generated from the words of the input sentence. An algorithm stored in the computer is applied to select which one of the number of general meaning nodes is intended by the user by comparing the input sentence pattern to the typical sentence patterns. The function associated with the selected general meaning node is executed.

ADVANTAGE - Allows knowledge engineer to build system that recognises any language or combination of language received from any source e.g keyboard or voice recognition.

Title Terms/Index Terms/Additional words: ANTICIPATE; MEANING; NATURAL; LANGUAGE; INTERFACE; SYSTEM; COMPUTER; APPLY; GENERAL; NUMBER; USER; INPUT; SENTENCE; STORAGE; NODE; ONE

Class Codes

International Classification (+ Attributes)

IPC + Level Value Position Status Version

G06F-0017/28 A I R 20060101

G06F-0017/28 C I R 20060101

US Classification, Issued: 395759

Alerting Abstract ...The method involves anticipating the general meaning of each of a number of likely user input sentences and storing in the computer a number of general meaning nodes, one for each anticipated...

Original Publication Data by Authority

Original Abstracts:

...it is abstracted by the system and compared to abstracted typical sentences in the knowledge base. This information, and other available information, is used by an algorithm to determine which of the general meaning nodes is intended...

12/3,K/1 (Item 1 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2008 European Patent Office. All rts. reserv.

01942987

TRANSLATION JUDGMENT DEVICE, METHOD, AND PROGRAM
TRANSLATIONSBEURTEILUNGSEINRICHTUNG, VERFAHREN UND PROGRAMM
DISPOSITIF, PROCEDE ET PROGRAMME D'EVALUATION DE TRADUCTION
PATENT ASSIGNEE:

Laboratory for Language Technology, (7037850), Incorporated 14-6-101
Hama-cho, Ashiya-shiHyogo 659-0025, (JP), (Applicant designated States:
all)

INVENTOR:

JACOBSON, YokoLab. for Language Technology Inc., 4-6-101, Hama-cho,
Ashiya-shi,Hyogo 6590025, (JP)

LEGAL REPRESENTATIVE:

Fuhlendorf, Jorn (3931), Patentanwalte Dreiss, Fuhlendorf, Steimle &
Becker, Postfach 10 37 62, 70032 Stuttgart, (DE)

PATENT (CC, No, Kind, Date): EP 1703419 A1 060920 (Basic)
WO 2005059771 050630

APPLICATION (CC, No, Date): EP 2004792480 041015; WO 2004JP15263 041015

PRIORITY (CC, No, Date): JP 2003416778 031215

DESIGNATED STATES: DE; ES; FR; GB; IT; NL

EXTENDED DESIGNATED STATES: AL; HR; LT; LV; MK

INTERNATIONAL PATENT CLASS (v7): G06F-017/28

INTERNATIONAL CLASSIFICATION (v8 + ATTRIBUTES):

IPC + Level Value Position Status Version Action Source Office:

G06F-0017/28 A I F B 20060101 20050707 H EP

ABSTRACT WORD COUNT: 242

NOTE:

Figure number on first page: 2

LANGUAGE (Publication,Procedural,Application): English; English; Japanese

FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200638	1767
SPEC A	(English)	200638	22204
Total word count - document A			23971
Total word count - document B			0
Total word count - documents A + B			23971

...SPECIFICATION in Step 108 are not counted twice even if the word
reappears in a natural sentence , in order to avoid repetitively
counting the coinciding words that appeared twice or more.

Thus even if the same coinciding word exists in multiple places of...

12/3,K/2 (Item 2 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2008 European Patent Office. All rts. reserv.

01784625

Document and pattern clustering method and apparatus
Dokument- und Mustergruppierungsverfahren und -Anordnung
Procede et dispositif de regroupement des documents et des formes
PATENT ASSIGNEE:

Hewlett-Packard Development Company, L.P., (4337790), 20555 S.H. 249,
Houston, TX 77070, (US), (Applicant designated States: all)

INVENTOR:

Kawatani, Takahiko, 1950-21-3-515,Mutsuura-cho, Kanazawa-Ku, Yokohama
Kanagawa 236-0032, (JP)

LEGAL REPRESENTATIVE:

Powell, Stephen David et al (52311), WILLIAMS POWELL Morley House 26-30
Holborn Viaduct, London EC1A 2BP, (GB)

PATENT (CC, No, Kind, Date): EP 1455285 A2 040908 (Basic)
EP 1455285 A2 040908
EP 1455285 A3 061220

APPLICATION (CC, No, Date): EP 2004251279 040305;

PRIORITY (CC, No, Date): JP 2003105867 030305; JP 200430629 040206

DESIGNATED STATES: AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR;

HU; IE; IT; LI; LU; MC; NL; PL; PT; RO; SE; SI; SK; TR
 EXTENDED DESIGNATED STATES: AL; LT; LV; MK
 INTERNATIONAL PATENT CLASS (V7): G06F-017/30
 INTERNATIONAL CLASSIFICATION (V8 + ATTRIBUTES):
 IPC + Level Value Position Status Version Action Source Office:
 G06F-0017/30 A I F B 20060101 20040705 H EP
 G06K-0009/62 A I L B 20060101 20061110 H EP
 ABSTRACT WORD COUNT: 112

NOTE:

Figure number on first page: 1

LANGUAGE (Publication,Procedural,Application): English; English; English
 FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200437	1874
SPEC A	(English)	200437	6522
Total word count - document A			8398
Total word count - document B			0
Total word count - documents A + B			8398

Document and pattern clustering method and apparatus

...ABSTRACT A3

In document (or pattern) clustering, the correct number of clusters and accurate assignment of each document (or pattern) to the correct cluster are attained. Documents (or patterns) describing the same topic (or object) are grouped, so a document (or pattern) group belonging to the same cluster has some commonality. Each topic (or object) has distinctive terms (or object features) or term (or object feature) pairs. When the closeness of each document (or pattern) to a given cluster is obtained, common information about the given cluster is extracted and used while the influence of terms (or object features) or term (or...

...SPECIFICATION M denote the number of kinds of the occurring terms, Dr)) denote the r-th document in a document set D consisting of R documents, Yr)) denote the number of sentences in document Dr)), and dry)) = (dry1)), ..., dryM))) T) denote...

...equation (1), the mn components of Sr) are given by
 Therefore, Sr)mm)) represents the number of sentences in which term m occur and Sr)mn)) represents the co - occurrence counts of sentences in which terms m and n co-occur. If each term does not occur twice or more in...

...U0) that stores the document frequencies of each term and each term in the input document set is obtained. Matrices U0)mm)) and U0)mn)) respectively denote the number of documents in...

12/3,K/3 (Item 3 from file: 348)
 DIALOG(R)File 348:EUROPEAN PATENTS
 (c) 2008 European Patent Office. All rts. reserv.

01274197

INFORMATION PROCESSING DEVICE AND INFORMATION PROCESSING METHOD, AND RECORDING MEDIUM

INFORMATIONSVERARBEITUNGSVORRICHTUNG UND INFORMATIONSVERARBEITUNGSVERFAHREN UND AUFNAHMEMEDIUM

PROCEDE ET DISPOSITIF INFORMATIQUE ET SUPPORT D'ENREGISTREMENT

PATENT ASSIGNEE:

Sony Corporation, (214028), 7-35, Kitashinagawa 6-chome, Shinagawa-ku, Tokyo 141-0001, (JP), (Applicant designated States: all)

INVENTOR:

IWAHASHI, Naoto Sony Computer Science Lab. Inc., 3-14-13, Higashi-Gotanda Shinagawa-ku, Tokyo 141-0022, (JP)

LEGAL REPRESENTATIVE:

Robinson, Nigel Alexander Julian (69551), D. Young & Co., 21 New Fetter Lane, London EC4A 1DA, (GB)

PATENT (CC, No, Kind, Date): EP 1146439 A1 011017 (Basic)

WO 200116794 010308

APPLICATION (CC, No, Date): EP 2000956860 000831; WO 2000JP5938 000831

PRIORITY (CC, No, Date): JP 99245461 990831
DESIGNATED STATES: DE; FR; GB; NL
EXTENDED DESIGNATED STATES: AL; LT; LV; MK; RO; SI
INTERNATIONAL PATENT CLASS (V7): G06F-017/28
ABSTRACT WORD COUNT: 104
NOTE:

Figure number on first page: 1

LANGUAGE (Publication,Procedural,Application): English; English; Japanese
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200142	995
SPEC A	(English)	200142	9049
Total word count - document A			10044
Total word count - document B			0
Total word count - documents A + B			10044

...SPECIFICATION analogousness. In addition, in the method using co-occurrence information, with respect to a large number of sentences, co-occurrence information of words appearing in those sentences are registered. Thus, the word analogousness is determined on the basis of...

12/3,K/4 (Item 4 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2008 European Patent Office. All rts. reserv.

01118380

METHOD AND SYSTEM FOR RETRIEVING RELEVANT DOCUMENTS FROM A DATABASE
METHODE UND VERFAHREN UM RELEVANTE DOKUMENTE IN EINER DATENBANK ZU FINDEN
PROCEDE ET SYSTEME POUR L'EXTRACTION DE DOCUMENTS PERTINENTS D'UNE BASE DE DONNEES

PATENT ASSIGNEE:

KCSL, Inc., (2910941), Suite 1012, 5160 Yonge Street, Toronto, Ontario
M2N 6L9, (CA), (Proprietor designated states: all)

INVENTOR:

KAUFMAN, Ilia, 18 Brandy Court, Toronto, Ontario M3B 3L3, (CA)

LEGAL REPRESENTATIVE:

Boyce, Conor et al (74271), F. R. Kelly & Co., 27 Clyde Road, Ballsbridge
, Dublin 4, (IE)

PATENT (CC, No, Kind, Date): EP 1086432 A1 010328 (Basic)

EP 1086432 B1 040407

WO 1999064964 991216

APPLICATION (CC, No, Date): EP 99924619 990607; WO 99CA531 990607

PRIORITY (CC, No, Date): US 88483 P 980608

DESIGNATED STATES: AT; BE; CH; CY; DE; DK; ES; FI; FR; GB; GR; IE; IT; LI;
LU; MC; NL; PT; SE

INTERNATIONAL PATENT CLASS (V7): G06F-017/30

NOTE:

No A-document published by EPO

LANGUAGE (Publication,Procedural,Application): English; English; English

FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS B	(English)	200415	779
CLAIMS B	(German)	200415	731
CLAIMS B	(French)	200415	857
SPEC B	(English)	200415	6447
Total word count - document A			0
Total word count - document B			8814
Total word count - documents A + B			8814

...SPECIFICATION sentence quantizer 60 then calculates the sum where the sum is over only those query- words that are present in the particular sentence Si)).

From these quantities, the sentence quantizer 60 calculates a position-independent sentence similarity using the following formula:
where #w(Si...

12/3,K/5 (Item 5 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS

(c) 2008 European Patent Office. All rts. reserv.

00395373

Method for determining the semantic relatedness of lexical items in a text.
Verfahren zur Bestimmung der semantischen Verwandtschaft zwischen lexikalischen
Einzelheiten in einem Text.

Methode pour determiner la parente semantique entre des unites lexicales
dans un texte.

PATENT ASSIGNEE:

BSO/BURO VOOR SYSTEEMONTWIKKELING B.V., (1192620), Kon. Wilhelminalaan 3,
P.O. Box 8348, NL-3503 RH Utrecht, (NL), (applicant designated states:
AT;BE;CH;DE;DK;ES;FR;GB;GR;IT;LI;LU;NL;SE)

INVENTOR:

Sadler, Victor, Livingstonelaan 304, NL-3526 HW Utrecht, (NL)

LEGAL REPRESENTATIVE:

de Bruijn, Leendert C. et al (19641), Nederlandsch Octrooibureau
Scheveningseweg 82 P.O. Box 29720, NL-2502 LS 's-Gravenhage, (NL)

PATENT (CC, No, Kind, Date): EP 386825 A1 900912 (Basic)

APPLICATION (CC, No, Date): EP 90200462 900226;

PRIORITY (CC, No, Date): NL 89587 890310

DESIGNATED STATES: AT; BE; CH; DE; DK; ES; FR; GB; GR; IT; LI; LU; NL; SE

INTERNATIONAL PATENT CLASS (V7): G06F-015/38;

ABSTRACT WORD COUNT: 217

LANGUAGE (Publication,Procedural,Application): English; English; English

FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	EPABF1	506
SPEC A	(English)	EPABF1	4217
Total word count - document A			4723
Total word count - document B			0
Total word count - documents A + B			4723

...ABSTRACT semantically related to each other, comprising the following
steps:

a) the retrieval from the said text corpus of a set of sentences
in which one or more of the given two or more lexical items...

...SPECIFICATION context makes it possible to find meaningful similarities
in the contextual patterns of semantically related words such as, in
the present example, the words DISCARD and REMOVE.

Even with the limited number of sentences used in this example, a
number of common contextual elements already appear. If the whole...

12/3,K/6 (Item 6 from file: 348)

DIALOG(R)File 348:EUROPEAN PATENTS

(c) 2008 European Patent Office. All rts. reserv.

00351453

Method and apparatus for producing an abstract of a document

Verfahren und Vorrichtung zur Herstellung einer Zusammenfassung eines
Dokumentes

Methode et dispositif pour produire un abrege d'un document

PATENT ASSIGNEE:

KABUSHIKI KAISHA TOSHIBA, (213130), 72, Horikawa-cho, Saiwai-ku,
Kawasaki-shi, Kanagawa-ken 210, (JP), (applicant designated states:
DE;FR;GB)

INVENTOR:

Doi, Miwako, (L-208) 30-1 Hisamoto Takatsu-ku, Kawasaki-shi Kanagawa-ken,
(JP)

LEGAL REPRESENTATIVE:

Lehn, Werner, Dipl.-Ing. et al (7471), Hoffmann Eitle, Patent- und
Rechtsanwälte, Postfach 81 04 20, 81904 München, (DE)

PATENT (CC, No, Kind, Date): EP 361464 A2 900404 (Basic)

EP 361464 A3 920902

EP 361464 B1 980812

APPLICATION (CC, No, Date): EP 89117915 890928;

PRIORITY (CC, No, Date): JP 88245967 880930

DESIGNATED STATES: DE; FR; GB

INTERNATIONAL PATENT CLASS (V7): G06F-017/24; G06F-017/30;

ABSTRACT WORD COUNT: 112

LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS B	(English)	9833	751
CLAIMS B	(German)	9833	714
CLAIMS B	(French)	9833	816
SPEC B	(English)	9833	3502
Total word count - document A			0
Total word count - document B			5783
Total word count - documents A + B			5783

...SPECIFICATION by this method. Moreover, the method has a drawback that, as the sentences with frequently appearing words are to be extracted, the number of sentences to be extracted also tends to become numerous, while a concise abstract is more desirable...

12/3,K/7 (Item 1 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

01630689 **Image available**

**MEDIA CONTENT ASSESSMENT AND CONTROL SYSTEMS
EVALUATION DE CONTENU MEDIA ET SYSTEMES DE GESTION**

Patent Applicant/Assignee:

WAGGENER EDSTROM WORLDWIDE INC, Three Centerpointe Drive Suite 300, Lake Oswego, Oregon 97035, US, US (Residence), US (Nationality), (For all designated states except: US)

Patent Applicant/Inventor:

GALLAGHER Daniel, 152 Se Spokane Street #6, Portland, Oregon 97202, US, US (Residence), US (Nationality), (Designated only for: US)

LIN Jia, 1545 Ne 96th Street, Seattle, Washington 98115, US, US

(Residence), CN (Nationality), (Designated only for: US)

STOFFREGEN Marc, 17217 Sw Sandhill Lane, Sherwood, Oregon 97140, US, US

(Residence), US (Nationality), (Designated only for: US)

Legal Representative:

BROOKS Michael Blaine (agent), 1445 East Los Angeles Ave Suite 301z, Simi Valley, California 93065, US

Patent and Priority Information (Country, Number, Date):

Patent: WO 200828070 A2 20080306 (WO 0828070)

Application: WO 2007US77286 20070830 (PCT/WO US2007077286)

Priority Application: US 2006824111 20060831; US 2007846866 20070829

Designated States:

(All protection types applied unless otherwise stated - for applications 2004+)

AE AG AL AM AT AU AZ BA BB BG BH BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM DO DZ EC EE EG ES FI GB GD GE GH GM GT HN HR HU ID IL IN IS JP KE KG KM KN KP KR KZ LA LC LK LR LS LT LU LY MA MD ME MG MK MN MW MX MY MZ NA NG NI NO NZ OM PG PH PL PT RO RS RU SC SD SE SG SK SL SM SV SY TJ TM TN TR TT TZ UA UG US UZ VC VN ZA ZM ZW

(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LT LU LV MC MT NL PL PT RO SE SI SK TR

(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG

(AP) BW GH GM KE LS MW MZ NA SD SL SZ TZ UG ZM ZW

(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext word Count: 10824

Fulltext Availability:

Detailed Description

Detailed Description

... word or phrase that is repeated. This example counts word frequency (WF) 4 11 for word occurrence, and for co - occurrence , counts sentence frequency (SF) 412 and paragraph frequency (PF) 413. Proximity counts, such as within three words , or phrase counts and co occurrences of phrase counts , e.g., sentence , paragraph, within specified word proximity, may also be included. The exemplary method may then rank...

12/3,K/8 (Item 2 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

01592878

**METHOD FOR SCORING CHANGES TO A WEBPAGE
PROCEDE POUR MARQUER LES CHANGEMENTS APPORTES A UNE PAGE WEB**

Patent Applicant/Assignee:

MONITOR110 INC, 58 East 11th Street, 3rd Floor, New York, New York 10003,
US, US (Residence), US (Nationality), (For all designated states
except: US)

Patent Applicant/Inventor:

STEWART Jeffrey A, 29 Great Jones, PHW, New York, New York 10012, US, US
(Residence), US (Nationality), (Designated only for: US)
AHMAD Shera, 98-01 67th Ave #9B, Rego Park, New York 11374, US, US
(Residence), US (Nationality), (Designated only for: US)

Legal Representative:

FERRARA Richard P (agent), Fish & Richardson P.C., P.O. Box 1022,
Minneapolis, Minnesota 55440-1022, US

Patent and Priority Information (Country, Number, Date):

Patent: WO 2007140364 A2 20071206 (WO 07140364)
Application: WO 2007US69880 20070529 (PCT/WO US2007069880)
Priority Application: US 2006808574 20060526; US 2007892945 20070305

Designated States:

(All protection types applied unless otherwise stated - for applications
2004+)

AE AG AL AM AT AU AZ BA BB BG BH BR BW BY BZ CA CH CN CO CR CU CZ DE DK
DM DO DZ EC EE EG ES FI GB GD GE GH GM GT HN HR HU ID IL IN IS JP KE KG
KM KN KP KR KZ LA LC LK LR LS LT LU LY MA MD ME MG MK MN MW MX MY MZ NA
NG NI NO NZ OM PG PH PL PT RO RS RU SC SD SE SG SK SL SM SV SY TJ TM TN
TR TT TZ UA UG US UZ VC VN ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LT LU LV MC MT
NL PL PT RO SE SI SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) BW GH GM KE LS MW MZ NA SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext word Count: 6224

Fulltext Availability:

Detailed Description

Detailed Description

... look to the text leading up to the third paragraph to see if any
predetermined **keywords** appear. The calculator may look to a preset
number of characters, **sentences**, paragraphs or the like leading to the
changed content to perform keyword analysis 430. For...

12/3,K/9 (Item 3 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

01509280 **Image available**

**METHOD FOR SEARCHING PATENT DOCUMENT BY APPLYING DEGREE OF SIMILARITY AND
SYSTEM THEREOF
PROCEDE DE RECHERCHE D'UN DOCUMENT DE BREVET PAR APPLICATION D'UN DEGRE DE
SIMILITUDE ET SYSTEME ASSOCIE**

Patent Applicant/Inventor:

KIM Jeong-Jin, 102-603 Cheonggu Apt., Hongje4-dong, Seodaemun-gu, Seoul
120-786, KR, KR (Residence), KR (Nationality), (Designated for all)

Legal Representative:

PARK Young-woo (agent), 5F., Seil Building, #727-13, Yoksam-dong,
Gangnam-gu, Seoul 135-921, KR

Patent and Priority Information (Country, Number, Date):

Patent: WO 200752883 A1 20070510 (WO 0752883)
Application: WO 2006KR3125 20060809 (PCT/WO KR2006003125)
Priority Application: KR 1020050104402 20051102

Designated States:

(All protection types applied unless otherwise stated - for applications

2004+)

AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM
DZ EC EE EG ES FI GB GD GE GH GM HN HR HU ID IL IN IS JP KE KG KM KN KP
KZ LA LC LK LR LS LT LU LV LY MA MD MG MK MN MW MX MZ NA NG NI NO NZ OM
PG PH PL PT RO RS RU SC SD SE SG SK SL SM SY TJ TM TN TR TT TZ UA UG US
UZ VC VN ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LT LU LV MC NL
PL PT RO SE SI SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) BW GH GM KE LS MW MZ NA SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: Korean

Fulltext word Count: 9176

Fulltext Availability:

Detailed Description

Detailed Description

... of "appearance frequency" of the patent document may be evaluated to be.

The "weight of **number** of **sentences** " indicates in how many sentences the **keywords** are found with respect to the **number** of **sentences** of the searched document. An amount of content of the searched document may be large...

...keywords appear in each of the documents, it is evaluated that the document in which **keywords** appear once in three sentences has a higher "weight of **number** of **sentences** ", as well as a higher weight of "appearance frequency".

As described above, when the additional...

...high when the keyword pair exists in the same sentence, and the higher the distance (**number** of **sentences**) between the **keywords** in the keyword pair is found to be, the lower the priority may become. In addition, the priority value may be...

12/3,K/10 (Item 4 from file: 349)

DIALOG(R)File 349:PCT FULLTEXT

(c) 2008 WIPO/Thomson. All rts. reserv.

01461711 **Image available**

GENERATING CHINESE LANGUAGE COUPLETS

GENERATION DE COUPLETS EN CHINOIS

Patent Applicant/Assignee:

MICROSOFT CORPORATION, One Microsoft way, Redmond, washington 98052-6399,
US, US (Residence), US (Nationality), (For all designated states
except: US)

Inventor(s):

ZHOU Ming, One Microsoft way, Redmond, washington 98052-6399, US,
(Designated for all)

SHUM Heung-Yeung, One Microsoft way, Redmond, washington 98052-6399, US,
(Designated for all)

Patent and Priority Information (Country, Number, Date):

Patent: WO 200705884 A2-A3 20070111 (WO 0705884)

Application: WO 2006US26064 20060703 (PCT/wo US2006026064)

Priority Application: US 2005173892 20050701

Designated States:

(All protection types applied unless otherwise stated - for applications

2004+)

AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM
DZ EC EE EG ES FI GB GD GE GH GM HN HR HU ID IL IN IS JP KE KG KM KN KP
KR KZ LA LC LK LR LS LT LU LV LY MA MD MG MK MN MW MX MZ NA NG NI NO NZ
OM PG PH PL PT RO RS RU SC SD SE SG SK SL SM SY TJ TM TN TR TT TZ UA UG
US UZ VC VN ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LT LU LV MC NL
PL PT RO SE SI SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) BW GH GM KE LS MW MZ NA SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English
Filing Language: English
Fulltext word Count: 6725

Fulltext Availability:
Detailed Description
Claims

Detailed Description

... be calculated using Equation 4 and $x=E1/81$, where E1 is the number of words appearing only once corresponding to b, and S is the total number of words in first scroll sentences of the training corpus corresponding to b, in the training data.
(2) For first scroll...

Claim

... the couplet corpus, wherein the sentence counts comprise number of sentences having a word x, number of sentences having a word y, and number of sentences having a co - occurrence of word x and word y.

5. The computer readable medium of claim 3, and further comprising constructing a Hidden...

12/3,K/11 (Item 5 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

01440335 **Image available**

COMPARING TEXT BASED DOCUMENTS COMPARAISON DE DOCUMENTS BASES SUR UN TEXTE

Patent Applicant/Assignee:

CURTIN UNIVERSITY OF TECHNOLOGY, Kent Street, Bentley, Western Australia 6102, AU, AU (Residence), AU (Nationality), (For all designated states except: US)

Patent Applicant/Inventor:

WILLIAMS Robert Francis, 5 Roche Court, Bull Creek, western Australia 6149, AU, AU (Residence), AU (Nationality),
DREHER Heinz, 195 Homestead Road, Mahogany Creek, western Australia 6072, AU, AU (Residence), AU (Nationality),

Legal Representative:

GRIFFITH HACK (agent), Level 19, 109 St. Georges Terrace, Perth, Western Australia 6000, AU

Patent and Priority Information (Country, Number, Date):

Patent: WO 2006119578 A1 20061116 (WO 06119578)

Application: WO 2006AU630 20060512 (PCT/WO AU2006000630)

Priority Application: AU 2005902424 20050513; AU 2005903032 20050610

Designated States:

(All protection types applied unless otherwise stated - for applications 2004+)

AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM
DZ EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KM KN KP KR
KZ LC LK LR LS LT LU LV LY MA MD MG MK MN MW MX MZ NA NG NI NO NZ OM PG
PH PL PT RO RU SC SD SE SG SK SL SM SY TJ TM TN TR TT TZ UA UG US UZ VC
VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LT LU LV MC NL
PL PT RO SE SI SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) BW GH GM KE LS MW MZ NA SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English
Filing Language: English
Fulltext word Count: 11078

Fulltext Availability:
Detailed Description

Detailed Description

... which words appear in the student essay; NoModelConcepts is the number of concepts for which words appear in the model essay; NoSentences is the number of sentences in the student essay; Nowords is the number

of words in the student essay; NonConceptualisedWordsRatio...

12/3,K/12 (Item 6 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

01072656 **Image available**

**WORD ASSOCIATION METHOD AND APPARATUS
PROCEDE ET APPAREIL D'ASSOCIATION DE MOTS**

Patent Applicant/Inventor:

ABIR Eli, 910 Route 35, Cross River, NY 10518, US, US (Residence), IL
(Nationality)

Legal Representative:

SONGER Michael J (et al) (agent), Arnold & Porter, 555 Twelfth Street,
NW, Washington, DC 20004-1206, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 2003102812 A1 20031211 (WO 03102812)

Application: WO 2003US2516 20030129 (PCT/WO US0302516)

Priority Application: US 2002157894 20020531; US 2002281997 20021029

Designated States:

(Protection type is "patent" unless otherwise stated - for applications
prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR
LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SC SD SE SG
SK SL TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT SE SI
SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext word Count: 27708

Fulltext Availability:

Detailed Description

English Abstract

...or near equivalent semantically. One method for associating words and
word strings includes querying a collection of documents with a
user-supplied word or word string input device 210), determining a
user-defined...

Detailed Description

... string).

47

Any combination of recurring patterns of words and word strings based on
the number of sentences in the database in which the word "Jets"
appears 3 words before "go to the game" when "tickets" appears 9 words
after "go to the game...

12/3,K/13 (Item 7 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

01063725 **Image available**

**ELECTRONIC DOCUMENT INDEXING SYSTEM AND METHOD
SYSTEME ET PROCEDE D'INDEXAGE DE DOCUMENTS ELECTRONIQUES**

Patent Applicant/Assignee:

HYPERBOLEX LIMITED, Level 2, 19 Tory Street, Wellington, NZ, NZ
(Residence), NZ (Nationality), (For all designated states except: US)

Patent Applicant/Inventor:

ANDERSON Roy Edward, 73 Donald Street, Karori, Wellington, NZ, NZ
(Residence), NZ (Nationality), (Designated only for: US)

Legal Representative:

ADAMS Matthew D (et al) (agent), A J Park, Huddart Parker Building, 6th
Floor, P.O. Box 949, Wellington 6015, NZ,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200394044 A1 20031113 (WO 0394044)
Application: WO 2003NZ82 20030505 (PCT/WO NZ0300082)
Priority Application: NZ 518744 20020503
Designated States:
(Protection type is "patent" unless otherwise stated - for applications prior to 2004)
AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR
LS LT LU LV MA MD MG MK MN MW MX MZ NI NO NZ OM PH PL PT RO RU SC SD SE
SG SK SL TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT RO SE
SI SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM
Publication Language: English
Filing Language: English
Fulltext word Count: 4674

Fulltext Availability:
Detailed Description

Detailed Description
... meet simple to complex lexical criteria including Boolean expressions.
A typical expression could be to "find all sentences having words
with the stem "weight" in combination with any of identify, count ,
sentence , document".

The collation of word use objects into a set of output sentences can be
...

12/3,K/14 (Item 8 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

01029401 **Image available**
CONTENT CONVERSION METHOD AND APPARATUS
PROCEDE ET APPAREIL DE CONVERSION DE CONTENU
Patent Applicant/Inventor:
ABIR Eli, 910 Route 35, Cross River, NY 10518, US, US (Residence), US
(Nationality)
Legal Representative:
SONGER Michael J (et al) (agent), Arnold & Porter, 555 Twelfth Street,
N.W., Washington, DC 20004-1206, US,
Patent and Priority Information (Country, Number, Date):
Patent: WO 200358374 A2-A3 20030717 (WO 0358374)

Application: WO 2002US29488 20020918 (PCT/WO US02029488)
Priority Application: US 200124473 20011221; US 2002157894 20020531
Designated States:
(Protection type is "patent" unless otherwise stated - for applications prior to 2004)
AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR
LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI
SK SL TJ TM TN TR TT TZ UA UG US UZ VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LU MC NL PT SE SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM
Publication Language: English
Filing Language: English
Fulltext word Count: 19291

Fulltext Availability:
Detailed Description

English Abstract
...determining the association between words in a language (Fig. 3). The
method includes providing a collection of documents (306), selecting
a first word or word strings, and a second word or word string...

French Abstract

...des associations entre des mots dans une langue. Le procede selon l'invention consiste a collecter des documents , et a choisir un premier mot ou suite de mots et un deuxieme mot ou...

Detailed Description

... word string).

Any combination of recurring patterns of words and word strings based on the number of sentences in the database in which the word "Jets" appears 3 words before "go to the game" when "tickets" appears 9 words after "go to the game...

12/3,K/15 (Item 9 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

00533612 **Image available**

METHOD AND SYSTEM FOR RETRIEVING RELEVANT DOCUMENTS FROM A DATABASE
PROCEDE ET SYSTEME POUR L'EXTRACTION DE DOCUMENTS PERTINENTS D'UNE BASE DE
DONNEES

Patent Applicant/Assignee:

KAUFMAN CONSULTING SERVICES LTD,
KAUFMAN Ilia,

Inventor(s):

KAUFMAN Ilia,

Patent and Priority Information (Country, Number, Date):

Patent: WO 9964964 A1 19991216

Application: WO 99CA531 19990607 (PCT/WO CA9900531)

Priority Application: US 9888483 19980608

Designated States:

(Protection type is "patent" unless otherwise stated - for applications prior to 2004)

AE AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB GD GE
GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MD MG MK
MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT UA UG US UZ VN
YU ZA ZW GH GM KE LS MW SD SL SZ UG ZW AM AZ BY KG KZ MD RU TJ TM AT BE
CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE BF BJ CF CG CI CM GA GN
GW ML MR NE SN TD TG

Publication Language: English

Fulltext word Count: 9941

Fulltext Availability:

Detailed Description

Detailed Description

... if t is a derivative query - word

where the sum is over only those query- words that are present in the particular sentence Si.

From these quantities , the sentence quantizer 60 calculates a position-independent sentence similarity using the following formula.

(4) Similar'tyd...

23/3,K/1 (Item 1 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2008 European Patent Office. All rts. reserv.

01624252

Summarization apparatus and method
Vorrichtung und Verfahren zur Zusammenfassung
Dispositif et procede pour faire des resumes

PATENT ASSIGNEE:

FUJITSU LIMITED, (211463), 1-1, Kamikodanaka 4-chome, Nakahara-ku,
Kawasaki-shi, Kanagawa 211-8588, (JP), (Applicant designated States:
all)

INVENTOR:

Nakao, Yoshio, Fujitsu Ltd., 4-1-1, Kamikodanaka, Nakahara-ku,
Kawasaki-shi, Kanagawa 211-8588, (JP)

LEGAL REPRESENTATIVE:

Stebbing, Timothy Charles et al (59643), Haseltine Lake, Imperial House,
15-19 Kingsway, London WC2B 6UD, (GB)

PATENT (CC, No, Kind, Date): EP 1338983 A2 030827 (Basic)
EP 1338983 A3 031217

APPLICATION (CC, No, Date): EP 2003008037 980116;

PRIORITY (CC, No, Date): JP 976777 970117

DESIGNATED STATES: DE; FR; GB

RELATED PARENT NUMBER(S) - PN (AN):

EP 855660 (EP 98300322)

INTERNATIONAL PATENT CLASS (V7): G06F-017/30

ABSTRACT WORD COUNT: 154

NOTE:

Figure number on first page: 2

LANGUAGE (Publication,Procedural,Application): English; English; English

FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200335	1478
SPEC A	(English)	200335	21974
Total word count - document A			23452
Total word count - document B			0
Total word count - documents A + B			23452

...SPECIFICATION of the document are output by adding a blank extraction
unit with its appearance position set at the end of the document in
step S73 and removing the unit in step S83. The description of the added
...

...is interested, etc. are stored in the user's preference 16. It also can
store keywords frequently appearing in such a document, the keywords
and question sentences often used in retrieval by a user, etc.
The user's knowledge 17 stores information...

...the users.

The document access log 18 accumulates the history of user's access to
documents and summaries.

The input document (group) 19 basically stores a document to be
summarized, and normally can be generated as any type of electronic
document. Practically...

23/3,K/3 (Item 3 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2008 European Patent Office. All rts. reserv.

02044267

A method and system for the orchestration of tasks on consumer electronics
Verfahren und System zum Steuern von Aufgaben in Unterhaltungselektronik
Procede et systeme pour orchestrer des taches en electronique de loisir

PATENT ASSIGNEE:

Samsung Electronics Co., Ltd., (7095030), 416 Maetan-Dong Yeongtong-Gu,
Suwon-si, Gyeonggi-Do, (KR), (Applicant designated States: all)

INVENTOR:

Messer, Alan, 225 Calle Marguerita, Los Gatos, California 95032, (US)
Kunjithapatham, Anugeetha, 243 Buena Vista Ave. Apt. 702,

Sunnyvale, California 94086, (US)
 LEGAL REPRESENTATIVE:
 Waddington, Richard et al (93232), Appleyard Lees, 15 Clare Road, Halifax
 HX1 2HY, (GB)
 PATENT (CC, No, Kind, Date): EP 1647884 A2 060419 (Basic)
 APPLICATION (CC, No, Date): EP 2005255590 050913;
 PRIORITY (CC, No, Date): US 948399 040922
 DESIGNATED STATES: AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR;
 HU; IE; IS; IT; LI; LT; LU; LV; MC; NL; PL; PT; RO; SE; SI; SK; TR
 EXTENDED DESIGNATED STATES: AL; BA; HR; MK; YU
 INTERNATIONAL CLASSIFICATION (V8 + ATTRIBUTES):
 IPC + Level Value Position Status Version Action Source Office:
 G06F-0009/44 A I F B 20060101 20060223 H EP
 ABSTRACT WORD COUNT: 149
 NOTE:
 Figure number on first page: 2

LANGUAGE (Publication,Procedural,Application): English; English; English
 FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200616	1703
SPEC A	(English)	200616	6691
Total word count - document A			8394
Total word count - document B			0
Total word count - documents A + B			8394

...SPECIFICATION selected/requested task suggestions.

For example, as noted task suggestions can be described as pseudo-sentences comprising a set of elements / terms that modify one another.

The present invention allows describing user tasks in an incremental and flexible way using pseudo-sentences which...

23/3,K/4 (Item 4 from file: 348)
 DIALOG(R)File 348:EUROPEAN PATENTS
 (c) 2008 European Patent Office. All rts. reserv.

02036183
 A method and system for presenting user tasks for the control of electronic devices
 Methode und Vorrichtung zur Darstellung von Benutzeranwendungsfälle zur Steuerung von elektronischen Geräten
 Methode et dispositif pour presenter des taches utilisateurs pour commander des appareils electroniques
 PATENT ASSIGNEE:
 Samsung Electronics Co., Ltd., (7095030), 416 Maetan-Dong Yeongtong-Gu, Suwon-si, Gyeonggi-Do, (KR), (Applicant designated States: all)
 INVENTOR:
 Messer, Alan, 225 Calle Marguerita, Los GatosCalifornia 95032, (US)
 Kunjithapatham, Anugeetha, 342 Buena Vista Avenue Apt., 702, SunnyvaleCalifornia 94086, (US)
 LEGAL REPRESENTATIVE:
 Waddington, Richard et al (93232), Appleyard Lees, 15 Clare Road, Halifax HX1 2HY, (GB)
 PATENT (CC, No, Kind, Date): EP 1640839 A1 060329 (Basic)
 APPLICATION (CC, No, Date): EP 2005255717 050915;
 PRIORITY (CC, No, Date): US 947774 040922
 DESIGNATED STATES: AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR; HU; IE; IS; IT; LI; LT; LU; LV; MC; NL; PL; PT; RO; SE; SI; SK; TR
 EXTENDED DESIGNATED STATES: AL; BA; HR; MK; YU
 INTERNATIONAL CLASSIFICATION (V8 + ATTRIBUTES):
 IPC + Level Value Position Status Version Action Source Office:
 G05B-0019/418 A I F B 20060101 20051119 H EP
 ABSTRACT WORD COUNT: 149
 NOTE:
 Figure number on first page: 2

LANGUAGE (Publication,Procedural,Application): English; English; English
 FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
----------------	----------	--------	------------

CLAIMS A	(English)	200613	1075
SPEC A	(English)	200613	6691
Total word count	- document A		7766
Total word count	- document B		0
Total word count	- documents A + B		7766

...SPECIFICATION selected/requested task suggestions.

For example, as noted task suggestions can be described as pseudo-sentences comprising a set of elements / terms that modify one another.

The present invention allows describing user tasks in an incremental and flexible way using pseudo-sentences which...

23/3,K/5 (Item 5 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2008 European Patent Office. All rts. reserv.

02036182

A method and system for describing consumer electronics using separate task and device descriptions

Methode und Vorrichtung zur Beschreibung von Haushaltselektronik unter Verwendung von separaten Aufgaben- und Gerätebeschreibungen

Methode et dispositif pour decire des produits electroniques en utilisant des descriptions de tache et des fonction separees

PATENT ASSIGNEE:

Samsung Electronics Co., Ltd., (7095030), 416 Maetan-Dong Yeongtong-Gu, Suwon-si, Gyeonggi-Do, (KR), (Proprietor designated states: all)

INVENTOR:

Messer, Alan, 225 Calle Marguerita, Los Gatos California 95032, (US)
Kunjithapatham, Anugeetha, 243 Buena Vista Ave., Apt., 702,, Sunnyvale, California 94086, (US)

LEGAL REPRESENTATIVE:

Waddington, Richard et al (93232), Appleyard Lees, 15 Clare Road, Halifax HX1 2HY, (GB)

PATENT (CC, No, Kind, Date): EP 1640838 A1 060329 (Basic)

EP 1640838 B1 071017

APPLICATION (CC, No, Date): EP 2005255716 050915;

PRIORITY (CC, No, Date): US 950121 040924

DESIGNATED STATES: AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR; HU; IE; IS; IT; LI; LT; LU; LV; MC; NL; PL; PT; RO; SE; SI; SK; TR

EXTENDED DESIGNATED STATES: AL; BA; HR; MK; YU

INTERNATIONAL CLASSIFICATION (V8 + ATTRIBUTES):

IPC + Level Value Position Status Version Action Source Office:

G05B-0019/418 A I F B 20060101 20051119 H EP

ABSTRACT WORD COUNT: 149

NOTE:

Figure number on first page: 2

LANGUAGE (Publication,Procedural,Application): English; English; English

FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
----------------	----------	--------	------------

CLAIMS A	(English)	200613	1073
----------	-----------	--------	------

CLAIMS B	(English)	200742	1125
----------	-----------	--------	------

CLAIMS B	(German)	200742	1058
----------	----------	--------	------

CLAIMS B	(French)	200742	1380
----------	----------	--------	------

SPEC A	(English)	200613	7030
--------	-----------	--------	------

SPEC B	(English)	200742	6847
--------	-----------	--------	------

Total word count	- document A		8104
------------------	--------------	--	------

Total word count	- document B		10410
------------------	--------------	--	-------

Total word count	- documents A + B		18514
------------------	-------------------	--	-------

...SPECIFICATION selected/requested task suggestions.

For example, as noted task suggestions can be described as pseudo-sentences comprising a set of elements / terms that modify one another.

The present invention allows describing user tasks in an incremental and flexible way using pseudo-sentences which...

...SPECIFICATION selected/requested task suggestions.

For example, as noted task suggestions can be described as pseudo-sentences comprising a set of elements / terms that modify one another.

The **present** invention allows describing user tasks in an incremental and flexible way using pseudo-sentences which...

23/3,K/6 (Item 6 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2008 European Patent Office. All rts. reserv.

01975121

Method and apparatus for classification of relative position of one or more text messages in an email thread

Methode und Apparat zur Klassifikation von relativen Positionen einer oder mehrerer Textnachrichten in einem Emailthread

Procede et dispositif pour classification de la position relative d'un ou plusieurs messages textes dans un thread des courriers electroniques

PATENT ASSIGNEE:

Avaya Technology Corp., (3148500), 211 Mount Airy Road, Basking Ridge, NJ 07920, (US), (Applicant designated States: all)

INVENTOR:

Bagga, Amit, 1054 Shadowlawn Drive, Green Brook, NJ 08812, (US)

Nenkova, Ani N., c/o Michele Banko, 302 18th Ave. East, Seattle, WA 98102, (US)

LEGAL REPRESENTATIVE:

Williams, David John et al (86433), Page White & Farrer Bedford House John Street, London, WC1N 2BF, (GB)

PATENT (CC, No, Kind, Date): EP 1591925 A2 051102 (Basic)

EP 1591925 A3 070620

APPLICATION (CC, No, Date): EP 2005251535 050315;

PRIORITY (CC, No, Date): US 833262 040427

DESIGNATED STATES: AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR; HU; IE; IS; IT; LI; LT; LU; MC; NL; PL; PT; RO; SE; SI; SK; TR

EXTENDED DESIGNATED STATES: AL; BA; HR; LV; MK; YU

INTERNATIONAL PATENT CLASS (V7): G06F-017/30

INTERNATIONAL CLASSIFICATION (V8 + ATTRIBUTES):

IPC + Level Value Position Status Version Action Source Office:

G06F-0017/30 A I F B 20060101 20050809 H EP

ABSTRACT WORD COUNT: 140

NOTE:

Figure number on first page: 3

LANGUAGE (Publication,Procedural,Application): English; English; English

FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200544	421
SPEC A	(English)	200544	3100
Total word count - document A			3522
Total word count - document B			0
Total word count - documents A + B			3522

...SPECIFICATION Generally, the present invention recognizes that emails that report problems or pose questions (most probably **root messages**) will be characterized by different punctuation than messages that contain answers or solutions.

v. Length of Email Message

The length of an email message, for example, in terms of the **number of sentences** can also be used as a feature. The length of an email message can be...

...root versus non-root word list 240 can be based on an examination of a **set** of root and non- **root messages** . Two dictionaries can be constructed with a first dictionary listing words typically occurring in non- **root messages** and another dictionary listing words typically occurring in **root messages** . The occurrence numbers can optionally be tested for statistical significance with the binomial test and...

...versus non-root classification task. In an exemplary implementation, the list of words typical for **root messages** was very short, while the list of words typical for non-root messages consisted of...

23/3,K/7 (Item 7 from file: 348)

DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2008 European Patent Office. All rts. reserv.

01241989

DOCUMENT SORTING METHOD, DOCUMENT SORTER, AND RECORDED MEDIUM ON WHICH
DOCUMENT SORTING PROGRAM IS RECORDED
DOKUMENTENSORTIERVERFAHREN, DOKUMENTENSORTIERER UND AUFZEICHNUNGSMEDIUM AUF
DEM DAS DOKUMENTENSORTIERPROGRAMM AUFGENOMMEN IST
PROCEDE DE TRI DE DOCUMENTS, APPAREIL DE TRI DE DOCUMENTS ET SUPPORT
ENREGISTRE SUR LEQUEL UN PROGRAMME DE TRI DE DOCUMENTS EST MEMORISE

PATENT ASSIGNEE:

SEIKO EPSON CORPORATION, (730004), 4-1, Nishishinjuku 2-chome,
Shinjuku-ku, Tokyo 163-0811, (JP), (Applicant designated States: all)

INVENTOR:

NAGAISHI, Michihiro, Seiko Epson Corporation, 3-5, Owa 3-chome, Suwa-shi,
Nagano 392-8502, (JP)

MIWA, Shinji, Seiko Epson Corporation, 3-5, Owa 3-ccome, Suwa-shi, Nagano
392-8502, (JP)

LEGAL REPRESENTATIVE:

Sturt, Clifford Mark et al (50502), Miller Sturt Kenyon 9 John Street,
London WC1N 2ES, (GB)

PATENT (CC, No, Kind, Date): EP 1124189 A1 010816 (Basic)

WO 200075810 001214

APPLICATION (CC, No, Date): EP 2000931690 000602; WO 2000JP3625 000602

PRIORITY (CC, No, Date): JP 99158498 990604; JP 99212501 990727

DESIGNATED STATES: DE; FR; GB

EXTENDED DESIGNATED STATES: AL; LT; LV; MK; RO; SI

INTERNATIONAL PATENT CLASS (V7): G06F-017/30

ABSTRACT WORD COUNT: 157

LANGUAGE (Publication,Procedural,Application): English; English; Japanese

FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
----------------	----------	--------	------------

CLAIMS A	(English)	200133	1230
----------	-----------	--------	------

SPEC A	(English)	200133	14072
--------	-----------	--------	-------

Total word count - document A	15302
-------------------------------	-------

Total word count - document B	0
-------------------------------	---

Total word count - documents A + B	15302
------------------------------------	-------

...SPECIFICATION title if such a part is detected. A third method is to
extract a predetermined **number** of **sentences** or words described at the
beginning of a **document** and employ the extracted sentence or words as
a title. The first, second, and third...displayed, and a categorization
result outputting unit 94 for outputting the categorization result
including the **cluster** -merging-process **information**.

The **clustering** unit 91 includes a **document** storage unit 911, a
sentence analyzer 912, a feature element extractor 913, a feature table

...

...title if such a part is detected. A third method is to extract a
predetermined **number** of **sentences** or words located at the **beginning**
of a **document** and employ the extracted sentence or words as a title.
The first, second, and third...

...of the feature table and categorizes the documents D1, D2,..., D7 into a
plurality of **clusters** according to semantic similarity. **Documents**
including a common feature element are detected on the basis of the
feature elements included...

23/3,K/8 (Item 8 from file: 348)

DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2008 European Patent Office. All rts. reserv.

01235196

Apparatus and method for generating a summary according to hierarchical
structure of topic

Gerat und Verfahren zum Erstellen einer der hierarchischen Struktur eines
Themas entsprechenden Zusammenfassung

Appareil et methode de generation d'un resume selon la structure
hierarchique du sujet

PATENT ASSIGNEE:

FUJITSU LIMITED, (211463), 1-1, Kamikodanaka 4-chome, Nakahara-ku,
 Kawasaki-shi, Kanagawa 211-8588, (JP), (Applicant designated States:
 all)
 INVENTOR:
 Nakao, Yoshio c/o Fujitsu Limited, 1-1, Kamikodanaka 4-chome Nakahara-ku,
 Kawasaki-shi, Kanagawa 211-8588, (JP)
 LEGAL REPRESENTATIVE:
 Mohun, Stephen John (76153), Haseltine Lake & Co., Imperial House, 15-19
 Kingsway, London WC2B 6UD, (GB)
 PATENT (CC, No, Kind, Date): EP 1071023 A2 010124 (Basic)
 EP 1071023 A3 021218
 APPLICATION (CC, No, Date): EP 2000305732 000707;
 PRIORITY (CC, No, Date): JP 99205061 990719
 DESIGNATED STATES: DE; GB
 EXTENDED DESIGNATED STATES: AL; LT; LV; MK; RO; SI
 INTERNATIONAL PATENT CLASS (v7): G06F-017/30
 ABSTRACT WORD COUNT: 84
 NOTE:
 Figure number on first page: 1

LANGUAGE (Publication,Procedural,Application): English; English; English
 FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200104	503
SPEC A	(English)	200104	14366
Total word count - document A			14869
Total word count - document B			0
Total word count - documents A + B			14869

...SPECIFICATION No. 7-36896 "Method and Apparatus for Generating Digest"
 extracts major expressions (word, etc.) as **seed** from a **document** based
 on the complexity of an expression (length of a word, etc.) and generates
 a...

...Application Laid-open No. 8-297677 "Method of Automatically Generating
 Digest of Topics" detects "topical **terms** " based on the **appearance**
 frequency of **words** in a document and generates a summary by extracting
sentences containing many major "topical terms".
 The second method judges the (relative) importance of sentences based
 ...

23/3,K/9 (Item 9 from file: 348)
 DIALOG(R)File 348:EUROPEAN PATENTS
 (c) 2008 European Patent Office. All rts. reserv.

01057144

INFORMATION MANAGEMENT AND KEY TERM RETRIEVAL
 INFORMATIONSMANAGEMENT UND WIEDERGEWINNUNG VON SCHLUSSELBEGRIFFEN
 GESTION D'INFORMATION ET EXTRACTION DES TERMES CLES
 PATENT ASSIGNEE:

BRITISH TELECOMMUNICATIONS public limited company, (846100), 81 Newgate
 Street, London EC1A 7AJ, (GB), (Proprietor designated states: all)
 INVENTOR:
 WEEKS, Richard, 44 Glemsford Close, Felixstowe, Suffolk IP11 8UG, (GB)
 LEGAL REPRESENTATIVE:
 Dutton, Erica L. G. et al (63161), BT Group Legal Services, Intellectual
 Property Department, 8th Floor, Holborn Centre 120 Holborn, London EC1N
 2TE, (GB)

PATENT (CC, No, Kind, Date): EP 1032896 A1 000906 (Basic)
 EP 1032896 B1 020327
 WO 9927469 990603
 APPLICATION (CC, No, Date): EP 98954628 981118; WO 98GB3468 981118
 PRIORITY (CC, No, Date): EP 97309446 971124
 DESIGNATED STATES: BE; CH; DE; ES; FI; FR; GB; IE; IT; LI; NL; SE
 INTERNATIONAL PATENT CLASS (v7): G06F-017/30

NOTE:

No A-document published by EPO
 LANGUAGE (Publication,Procedural,Application): English; English; English
 FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS B	(English)	200213	776

CLAIMS B	(German)	200213	763
CLAIMS B	(French)	200213	833
SPEC B	(English)	200213	10645
Total word count	- document A		0
Total word count	- document B		13017
Total word count	- documents A + B		13017

...SPECIFICATION Figures 5, 6 and 7. within the inner scanning loop, having selected a particular word **group element** WG(S,k,i), and having established that it is not null, remaining word groups...

...those with a higher value of k within the same sentence and those in later **sentences** only, are checked for matching word groups. For **each match found**, the weighting of word group WG(S,k,i) is incremented and the matching word group is set to...

23/3,K/10 (Item 10 from file: 348)
 DIALOG(R)File 348:EUROPEAN PATENTS
 (c) 2008 European Patent Office. All rts. reserv.

00672789

Dictionary creation supporting system
 Unterstützungssystem zur Herstellung von Wörterbüchern
 Systeme de support pour la creation de dictionnaires

PATENT ASSIGNEE:

KABUSHIKI KAISHA TOSHIBA, (213130), 72, Horikawa-cho, Saiwai-ku,
 Kawasaki-shi, Kanagawa-ken 210-8572, (JP), (Proprietor designated
 states: all)

INVENTOR:

Hirakawa, Hideki, 1-18-24, Katida-Minami, Kohoku-ku, Yokohama-shi,
 Kanagawa-ken, (JP)
 Kumano, Akira, 7-4-401, Nakadai, Higashiterao, Tsurumi-ku, Yokohama-shi,
 Kanagawa-ken, (JP)

LEGAL REPRESENTATIVE:

Lehn, Werner, Dipl.-Ing. et al (7474), Hoffmann Eitle, Patent- und
 Rechtsanwälte, Arabellastrasse 4, 81925 München, (DE)

PATENT (CC, No, Kind, Date): EP 645720 A2 950329 (Basic)
 EP 645720 A3 951129
 EP 645720 B1 010801

APPLICATION (CC, No, Date): EP 94114789 940920;

PRIORITY (CC, No, Date): JP 93232649 930920

DESIGNATED STATES: DE; FR; GB

INTERNATIONAL PATENT CLASS (V7): G06F-017/27; G06F-017/28

ABSTRACT WORD COUNT: 164

NOTE:

Figure number on first page: 1

LANGUAGE (Publication,Procedural,Application): English; English; English

FULLTEXT AVAILABILITY:

Available Text	Language	Update	word Count
CLAIMS A	(English)	EPAB95	1569
CLAIMS B	(English)	200131	1485
CLAIMS B	(German)	200131	1555
CLAIMS B	(French)	200131	1829
SPEC A	(English)	EPAB95	9103
SPEC B	(English)	200131	9012
Total word count	- document A		10674
Total word count	- document B		13881
Total word count	- documents A + B		24555

...SPECIFICATION TEJUN" (Japanese word generically meaning operation procedure) are outputted in this order as these composite **words appear** in this order in the original **sentences**. In this case, the operation of the registration word selection processing s1 proceeds as follows...

...and the value of its superficial position (a value of "mds") is the smallest. This **element** is set as the **element 1**.

(2) The element 1 is deleted from the output information source file.

(3) The **element 1**. These **elements** are set as the **element 2**, (sup(...), **element N**.

(4) The element 2, (sup(...)) , element N are deleted from the output information source...

...SPECIFICATION TEJUN" (Japanese word generically meaning operation procedure) are outputted in this order as these composite words appear in this order in the original sentences . In this case, the operation of the registration word selection processing sl proceeds as follows...

...and the value of its superficial position (a value of "mds") is the smallest. This element is set as the element 1.

(2) The element 1 is deleted from the output information source file.

(3) The...

...is searched to take out the elements having the same registration knowledge information as the element 1. These elements are set as the element 2, (center dot)(center dot)(center dot)(center dot)(center dot)(center dot)(center dot)(center dot)...

23/3,K/11 (Item 11 from file: 348)
 DIALOG(R)File 348:EUROPEAN PATENTS
 (c) 2008 European Patent Office. All rts. reserv.

00656039

A method and system of information retrieval
 Ein Verfahren und System zur Informationswiederauffindung
 Un procede et systeme pour le recouvrement d'informations

PATENT ASSIGNEE:

XEROX CORPORATION, (219783), Xerox Square, Rochester, New York 14644,
 (US), (Proprietor designated states: all)

INVENTOR:

Kupiec, Julian M., 10070 Craft Drive, Cupertino, California 95014, (US)

LEGAL REPRESENTATIVE:

Grunecker, Kinkeldey, Stockmair & Schwanhauser Anwaltssozietat (100721)
 , Maximilianstrasse 58, 80538 Munchen, (DE)

PATENT (CC, No, Kind, Date): EP 631244 A2 941228 (Basic)
 EP 631244 A3 950222
 EP 631244 B1 021106

APPLICATION (CC, No, Date): EP 94302927 940425;

PRIORITY (CC, No, Date): US 82938 930624

DESIGNATED STATES: DE; FR; GB

INTERNATIONAL PATENT CLASS (V7): G06F-017/30

ABSTRACT WORD COUNT: 125

NOTE:

Figure number on first page: 1

LANGUAGE (Publication,Procedural,Application): English; English; English

FULLTEXT AVAILABILITY:

Available Text	Language	Update	word Count
CLAIMS A	(English)	EPABF2	913
CLAIMS B	(English)	200245	861
CLAIMS B	(German)	200245	856
CLAIMS B	(French)	200245	979
SPEC A	(English)	EPABF2	20043
SPEC B	(English)	200245	20109
Total word count - document A			20960
Total word count - document B			22805
Total word count - documents A + B			43765

...ABSTRACT A2

A computerized method for organizing information retrieval based on the content of a set of primary documents . The method generates answer hypotheses based on text found in the primary documents and, typically, a natural-language input string such as a question. The answer hypotheses can...

...A text corpus (12) can be queried to provide verification evidence not present in the primary documents . In another aspect the method is implemented in the context of a larger two-phase...

...SPECIFICATION is substituted for a placeholder or placeholders in the template.

6.5 Matching Templates Against Primary Documents

In step 264 an attempt is made to verify the linguistic relation under consideration for the hypothesis under consideration in the context of the **primary documents**. This is done by matching the filled-in templates generated in step 263 against the **primary documents**. In other words, sentences in which the hypothesis appears in the context of a template are sought in the **primary documents**. Any such sentences found are retained in association with the hypothesis as verification evidence for...

...SPECIFICATION is substituted for a placeholder or placeholders in the template.

6.5 Matching Templates Against Primary Documents

In step 264 an attempt is made to verify the linguistic relation under consideration for the hypothesis under consideration in the context of the **primary documents**. This is done by matching the filled-in templates generated in step 263 against the **primary documents**. In other words, sentences in which the hypothesis appears in the context of a template are sought in the **primary documents**. Any such sentences found are retained in association with the hypothesis as verification evidence for...

23/3,K/12 (Item 12 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2008 European Patent Office. All rts. reserv.

00522837

Method and system for natural language translation

Verfahren und System zur Sprachubersetzung

Methode et systeme de traduction en langage naturel

PATENT ASSIGNEE:

International Business Machines Corporation, (200120), Old Orchard Road,
Armonk, N.Y. 10504, (US), (Proprietor designated states: all)

INVENTOR:

Brown, Peter Fitzhugh, 390 Riverside Drive, Apt. 14A, New York, New York
10025, (US)

Cocke, John, 87 Pound Ridge Road, Bedford, New York 10506, (US)

Della Pietra, Stephen Andrew, 113 Meyer Oval, Pearl River, New York 10965
(US)

Della Pietra, Vincent Joseph, 129 Sunset Road, Blauvelt, New York 10913,
(US)

Jelinek, Frederick, 511 Scarborough Road, Briarcliff Manor, New York
10510, (US)

Lai, Jennifer Ceil, Cat Rock Road, Garrison, New York 10524, (US)

Mercer, Robert Leroy, 669 Viewland Drive, Yorktown Heights, New York
10598, (US)

LEGAL REPRESENTATIVE:

Teufel, Fritz, Dipl.-Phys. et al (11855), IBM Deutschland

Informationssysteme GmbH, Patentwesen und Urheberrecht, 70548 Stuttgart
, (DE)

PATENT (CC, No, Kind, Date): EP 525470 A2 930203 (Basic)

EP 525470 A3 940413

EP 525470 B1 000405

APPLICATION (CC, No, Date): EP 92111725 920710;

PRIORITY (CC, No, Date): US 736278 910725

DESIGNATED STATES: DE; FR; GB; IT

INTERNATIONAL PATENT CLASS (V7): G06F-017/28

ABSTRACT WORD COUNT: 244

NOTE:

Figure number on first page: 1

LANGUAGE (Publication,Procedural,Application): English; English; English

FULLTEXT AVAILABILITY:

Available Text	Language	Update	word Count
----------------	----------	--------	------------

CLAIMS B	(English)	200014	1928
----------	-----------	--------	------

CLAIMS B	(German)	200014	1620
----------	----------	--------	------

CLAIMS B	(French)	200014	2129
----------	----------	--------	------

SPEC B	(English)	200014	39424
--------	-----------	--------	-------

Total word count - document A	0
-------------------------------	---

Total word count - document B	45101
-------------------------------	-------

Total word count - documents A + B	45101
------------------------------------	-------

...ABSTRACT model are combined into a combined score for each intermediate target-structure hypothesis. Finally, a set of target- text hypotheses is produced by transducing the highest scoring target-structure hypotheses into portions of text...

...SPECIFICATION from several years of the proceedings of the Canadian parliament. From these translations, a training data set is chosen comprising those pairs for which both the English sentence and the French ...

...that abound in the text, a English vocabulary is chosen consisting of all of those words that appear at least twice in English sentences in the data, and as a French vocabulary is chosen consisting of all those words that appear at least twice in French sentences in the data. All other words are replaced with a special unknown English word or unknown...

23/3,K/13 (Item 13 from file: 349)

DIALOG(R)File 349:PCT FULLTEXT

(c) 2008 WIPO/Thomson. All rts. reserv.

01191853 **Image available**

METHOD AND SYSTEM FOR ENHANCED DATA SEARCHING

PROCEDE ET SYSTEME POUR UNE RECHERCHE AMELIOREE DE DONNEES

Patent Applicant/Assignee:

INSIGHTFUL CORPORATION, 1700 Westlake Avenue North, Suite 500, Seattle, WA 98109-3044, US, US (Residence), US (Nationality), (For all designated states except: US)

Patent Applicant/Inventor:

MARCHISIO Giovanni B, 9815 Northeast 130th Place, Unit 303, Kirkland, WA 98034, US, US (Residence), US (Nationality), (Designated only for: US)
KOPERSKI Krzysztof, 2311 Yale Avenue East, Apt. D., Seattle, WA 98102, US, US (Residence), CA (Nationality), (Designated only for: US)
LIANG Jisheng, 6343 - 114th Avenue Southeast, Bellevue, WA 98006, US, US (Residence), CN (Nationality), (Designated only for: US)
MURUA Alejandro, 1310 East Thomas Street, Apt. 302, Seattle, WA 98102, US, US (Residence), CL (Nationality), (Designated only for: US)
NGUYEN Thien, 22220 - 98th Avenue West, Edmonds, WA 98020, US, US (Residence), US (Nationality), (Designated only for: US)
TUSK Carsten, 20912 4th Avenue South, Seattle, WA 98198, US, US (Residence), DE (Nationality), (Designated only for: US)
DHILLON Navdeep S, 8011 29th Avenue NW, Seattle, WA 98117, US, US (Residence), US (Nationality), (Designated only for: US)
POCHMAN Lubos, P.O. Box 3807, Breckenridge, CO 80424, US, US (Residence), US (Nationality), (Designated only for: US)

Legal Representative:

BIERMAN Ellen M (et al) (agent), Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 2004114163 A2-A3 20041229 (WO 04114163)

Application: WO 2004US4099 20040212 (PCT/WO US04004099)

Priority Application: US 2003371399 20030219

Designated States:

(All protection types applied unless otherwise stated - for applications 2004+)

AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM
DZ EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC
LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NA NI NO NZ OM PG PH PL PT RO
RU SC SD SE SG SK SL SY TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT RO SE
SI SK TR

(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG

(AP) BW GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW

(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext word Count: 31623

Patent and Priority Information (Country, Number, Date):

Patent: ... 20041229

Fulltext Availability:
Detailed Description

English Abstract

Methods and systems for syntactically indexing and searching **data sets** to achieve more accurate search results and for indexing and searching **data sets** using entity tags alone or in combination therewith are provided. Example embodiments provide a Syntactic Query Engine ("SQE") that parses, indexes, and stores a **data set**, as well as processes natural language queries subsequently submitted against the **data set**. The SQE comprises a Query Preprocessor, a **Data Set** Preprocessor, a Query Builder, a **Data Set** Indexer, an Enhanced Natural Language Parser ("ENLP"), a **data set** repository, and, in some embodiments, a user interface. After preprocessing the **data set**, the SQE parses the **data set** according to a variety of levels of parsing and determines as appropriate the entity tags...
...grammatical roles of each term to generate enhanced data representations for each object in the **data set**. The SQE indexes and stores these enhanced **data** representations in the **data set** repository. Upon subsequently receiving a query, the SQE parses the query also using a variety of parsing levels and searches the indexed stored **data set** to locate **data** that contains similar terms used in similar grammatical roles and/or with similar entity tag...

Publication Year: 2004

Detailed Description

... subject/ verb/preposition/verb modifier/object; and
noun/noun modifier.

24

Such support includes locating **sentences** in which the designated **terms appear** in the associated designated syntactic or grammatical role, as well as locating, when contextually appropriate...

23/3,K/14 (Item 14 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

01123033

METHOD AND SYSTEM FOR USING QUERY INFORMATION TO ENHANCE CATEGORIZATION AND NAVIGATION WITHIN THE WHOLE KNOWLEDGE BASE
PROCEDE ET SYSTEME PERMETTANT D'UTILISER DES INFORMATIONS DE REQUETES POUR AMELIORER LA CATEGORISATION ET LA NAVIGATION DANS LA TOTALITE DE LA BASE DE CONNAISSANCES

Patent Applicant/Assignee:

KENNETH Nadav, 30 Ha-Mazbiim Street, 69935 Tel Aviv, IL, IL (Residence),
IL (Nationality), (For all designated states except: US)
MIZRAHI Moshe, 21 Avner Street, 69937 Tel Aviv, IL, IL (Residence), IL
(Nationality), (For all designated states except: US)

Patent Applicant/Inventor:

SEBBANE Danny, 18 Adam Hacohen Street, 64585 Tel Aviv, IL, IL (Residence),
IL (Nationality)

Legal Representative:

NAOMI ASSIA LAW OFFICES (agent), 32 Habarzel Street, Ramat Hachaya1,
69710 Tel Aviv, IL,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200444896 A2-A3 20040527 (WO 0444896)

Application: WO 2003IL938 20031110 (PCT/WO IL03000938)

Priority Application: US 2002425728 20021113

Designated States:

(Protection type is "patent" unless otherwise stated - for applications prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM
DZ EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC
LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NI NO NZ OM PG PH PL PT RO RU
SC SD SE SG SK SL SY TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT RO SE
SI SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English
Filing Language: English
Fulltext word Count: 5299

Patent and Priority Information (Country, Number, Date):

Patent: ... 20040527
Fulltext Availability:
Detailed Description

English Abstract

...is disclosed to create some structure from the knowledge base of an organization, the knowledge base including a document database (DB) and queries submitted by users concerning the documents, wherein the method performs monitoring...
Publication Year: 2004

Detailed Description

... to each other, as described in step 4 below. Queries are associated with phrases (or sentences) and clusters are associated with documents. Thus, words that appear in queries have an added component relative to those that only appear in documents. A...
...documents; and phrases. A word that also appears in queries has a 4-dimensioal vector: documents; phrases; clusters; and queries. A vector is used to represent the distribution of the word in the...

23/3,K/15 (Item 15 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

01119799 **Image available**

POST-PROCESSING SYSTEM AND METHOD FOR CORRECTING MACHINE RECOGNIZED TEXT
SYSTEME DE POST-TRAITEMENT POUR LA CORRECTION DE TEXTES LISIBLES PAR
MACHINE

Patent Applicant/Assignee:

MATSUSHITA ELECTRIC INDUSTRIAL CO LTD, Matsushita IMP Bldg., 19F, 1-3-7,
Shiomi, Shuo-ku, Osaka 540-6319, JP, JP (Residence), JP (Nationality)

Inventor(s):

MA Yue, 6 Tiffany Court, West Windsor, NJ 08550, US,
GUO Jinhong Katherine, 6 Tiffany Court, West Windsor, NJ 08850, US,
LI Mu, 5/F, No. 49 Zhichun Road, Beijing, 100080, CN,
TONG Yu-kun, 226, Sixth Student Dorm, Northeastern University, Shenyang,
110004, CN,
YAO Tian-shun, 13-1, 19 Building wanghubeilu, Shenyang, 110004, CN,
ZHU Jing-bo, 5-1-3, Qingniangongyu, Northeastern University, Sheyang,
110004, CN,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200442641 A2-A3 20040521 (WO 0442641)
Application: WO 2003IB6487 20031104 (PCT/WO IB03006487)
Priority Application: US 2002288645 20021104

Designated States:

(Protection type is "patent" unless otherwise stated - for applications prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM
DZ EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC
LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NI NO NZ OM PG PH PL PT RO RU
SC SD SE SG SK SL SY TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT RO SE
SI SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) BW GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English
Filing Language: English
Fulltext word Count: 10790

Patent and Priority Information (Country, Number, Date):

Patent: ... 20040521
Fulltext Availability:
Detailed Description

English Abstract

...OCR) engine and apparatus to perform the method. This exemplary method includes segmenting the character **data** into a **set** of initial words. The set of initial words is word level processed to determine at...
Publication Year: **2004**

Detailed Description

... no final sentence is selected, but the candidate word sets are examined and any candidate **words** that do not **appear** in at least one of the candidate **sentences** having the highest POS tri-gram cost are removed. If only one candidate word remains...

...no final sentence is selected, but the candidate word sets are examined and any candidate **words** that do not **appear** in at least one of the candidate **sentences** having the highest word tri-gram cost are removed. If only one candidate word remains...

23/3,K/16 (Item 16 from file: 349)

DIALOG(R)File 349:PCT FULLTEXT

(c) 2008 WIPO/Thomson. All rts. reserv.

01013799 **Image available**

METHOD AND APPARATUS FOR TEXTUAL EXPLORATION DISCOVERY

PROCEDE ET APPAREIL D'EXPLORATION ET DE DECOUVERTE TEXTUELLE

Patent Applicant/Assignee:

FORINNOVA AS, Thormohlensgate 55, N-5008 Bergen, NO, NO (Residence), NO (Nationality), (For all designated states except: US)

Patent Applicant/Inventor:

AARSKOG Brit Helle, Sore Furudalen 2, N-5098 Bergen, NO, NO (Residence), NO (Nationality), (Designated only for: US)

Legal Representative:

AS BERGEN PATENTKONTOR (agent), P.O. Box 1998, Nordnes, N-5817 Bergen, NO

Patent and Priority Information (Country, Number, Date):

Patent: WO 200342859 A2-A3 **20030522** (WO 0342859)

Application: WO 2002NO423 20021115 (PCT/WO NO0200423)

Priority Application: NO 20015581 20011115

Designated States:

(Protection type is "patent" unless otherwise stated - for applications prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR
LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI
SK SL TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LU MC NL PT SE SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext word Count: 68587

Patent and Priority Information (Country, Number, Date):

Patent: ... **20030522**

Fulltext Availability:

Detailed Description

Publication Year: **2003**

Detailed Description

... in 'meaning' is ignored when the goal is to locate or detect a relative small **set** of **documents** out of a **collection** comprising millions of **documents**. In the present invention the goal is quite different - the variations in wordings are captured...aiming at directing the users' attention to zones where the link set for the constituent **sentences** indicate a bundle of focused **words** or several **co - occurring** focused **words**. More specifically, the idea is: when the user selects a document for exploration, a text...otherwise similar sentences can be notified as different.

The following elements constitute parts of the **information** in the link **sets** and the **words** listed in order of **appearance** in the **sentences**

. Sentences marked as 1 and 2 share 4 noun elements, of the 4 noun elements are...

23/3,K/17 (Item 17 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

00987311 **Image available**

METHOD AND SYSTEM FOR ENHANCED DATA SEARCHING
PROCEDE ET SYSTEME PERMETTANT D'EFFECTUER UNE RECHERCHE AMELIOREE DES
DONNEES

Patent Applicant/Assignee:

INSIGHTFUL CORPORATION, Suite 500, 1700 Westlake Avenue North, Seattle,
WA 98109-3044, US, US (Residence), US (Nationality), (For all
designated states except: US)

Patent Applicant/Inventor:

MARCHISIO Giovanni B, Unit 303, 9815 NE 130th Place, Kirkland, WA 98034,
US, US (Residence), US (Nationality), (Designated only for: US)
KOPERSKI Krzysztof, Apt. D, 2311 Yale Avenue East, Seattle, WA 98102, US,
US (Residence), CA (Nationality), (Designated only for: US)
LIANG Jisheng, 6343 114th Avenue Southeast, Bellevue, WA 98006, US, US
(Residence), CN (Nationality), (Designated only for: US)
MURUA Alejandro, Apt. 302, 1310 East Thomas Street, Seattle, WA 98102, US
, US (Residence), CL (Nationality), (Designated only for: US)
NGUYEN Thien, 22220 98th Avenue West, Edmonds, WA 98020, US, US
(Residence), US (Nationality), (Designated only for: US)

Legal Representative:

BIERMAN Ellen M (et al) (agent), Seed Intellectual Property Law Group
PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200317143 A2-A3 20030227 (WO 0317143)
Application: WO 2002US25756 20020814 (PCT/WO 0225756)
Priority Application: US 2001312385 20010814; US 20017299 20011108

Designated States:

(Protection type is "patent" unless otherwise stated - for applications
prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR
LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI
SK SL TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LU MC NL PT SE SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext word Count: 35592

Patent and Priority Information (Country, Number, Date):

Patent: ... 20030227

Fulltext Availability:

Detailed Description

English Abstract

Methods and systems for syntactically indexing and searching data sets
to achieve more accurate search results are provided. Example
embodiments provide a Syntactic Query Engine ("SQE") that parses,
indexes, and stores a data set, as well as processes natural language
queries subsequently submitted against the data set. The SQE
comprises a Query Preprocessor, a Data Set Preprocessor, a Query
Builder, a Data Set Indexer, an Enhanced Natural Language Parser
("ENLP"), a data set repository, and, in some embodiments, a user
interface. After preprocessing the data set, the SQE parses the data
set and determines the syntactic and grammatical roles of each term to
generate enhanced data representations for each object in the data set.
The SQE indexes and stores these enhanced data representations in the
data set repository. Upon subsequently receiving a query, the SQE
parses the query similarly and searches the indexed stored data set
to locate data that contains similar terms used in similar grammatical
roles. In this manner, the SQE is...

Publication Year: 2003

Detailed Description

... modifier;
subject/ verb/preposition/verb modifier/object; and
noun/noun modifier.

Such support includes locating **sentences** in which the designated **terms appear** in the associated designated syntactic or grammatical role, as well as locating, when contextually appropriate, **sentences** in which the designated **terms appear** but where the designated roles ...may be implemented to recognize any number of programmable attributes in natural language queries and **data sets** (described in detail as "preferences" with reference to Figure 15). In one embodiment, these attributes...

23/3,K/18 (Item 18 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

00982578 **Image available**

COMPUTER BASED SUMMARIZATION OF NATURAL LANGUAGE DOCUMENTS
RECAPITULATION INFORMATIQUE DE DOCUMENTS EN LANGAGE NATUREL

Patent Applicant/Assignee:

INVENTION MACHINE CORPORATION, 133 Portland Street, Boston, MA 02114-1722
, US, US (Residence), US (Nationality), (For all designated states
except: US)

Patent Applicant/Inventor:

BATCHILO Leonid, 35 Moraine Street, Belmont, MA 02478, US, US (Residence)
, BY (Nationality), (Designated only for: US)

TSOURIKOV Valery, 177 Marlborough Street, Apt. 8, Boston, MA 02116, US,
US (Residence), BY (Nationality), (Designated only for: US)

SOVPEL Igor, 3/1 Voronyanskogo Street, Apt. 193, Minsk, 220029, BY, BY
(Residence), BY (Nationality), (Designated only for: US)

Legal Representative:

MELLO David M (agent), McDermott, Will & Emery, 28 State Street, Boston,
MA 02109, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200312661 A1 **20030213** (WO 0312661)

Application: WO 2002US24259 20020731 (PCT/WO US0224259)

Priority Application: US 2001308886 20010731

Designated States:

(Protection type is "patent" unless otherwise stated - for applications
prior to 2004)

AE AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB GD GE
GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MD MG MK
MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT UA UG US UZ VN
YU ZA ZW

(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LU MC NL PT SE SK TR

(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG

(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW

(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext word Count: 9334

Patent and Priority Information (Country, Number, Date):

Patent: ... **20030213**

Fulltext Availability:

Detailed Description

Publication Year: **2003**

Detailed Description

... conduct elementary morphological words analysis. Commonly, the summary was made up from the sentences of **initial text** that received the highest rank, or that met some other criteria. The statistics, in such cases, were **collected** on **text** word usage rate. That is, the more the word was found in the text, the weightier it was considered. Auxiliary words and other of a word in a **document set** was taken into consideration. Such estimation is discussed in U.S.

Patent No. 6, 128...values, such as the average number of words and

symbols in a sentence, the average number of sentences in the paragraph, and so on. Then, the topic, of the document is defined on...

23/3,K/19 (Item 19 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

00895471 **Image available**

METHOD AND APPARATUS FOR TRANSFORMING CONTENTS ON THE WEB
PROCEDE ET APPAREIL PERMETTANT LA TRANSFORMATION DE CONTENUS EN LIGNE
Patent Applicant/Assignee:

HEWLETT-PACKARD COMPANY, Intellectual Property Administration, P.O. Box
272400, Fort Collins, CO 80527-2400, US, US (Residence), US
(Nationality), (For all designated states except: US)

Patent Applicant/Inventor:

YAMAMOTO Akio, 33-12-409, Sakuradai, Aoba-ku, Yokohama-shi, Kanagawa, JP,
JP (Residence), JP (Nationality), (Designated only for: US)

Legal Representative:

GREELEY Paul D (agent), Ohlandt, Greeley, Ruggiero & Perle, L.L.P., 10th
Floor, One Landmark Square, Stamford, CT 06901-2682, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200229590 A1 20020411 (WO 0229590)

Application: WO 2001US30691 20011002 (PCT/WO US0130691)

Priority Application: JP 2000302728 20001002

Designated States:

(Protection type is "patent" unless otherwise stated - for applications
prior to 2004)

CN KR US

(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR

Publication Language: English

Filing Language: English

Fulltext word Count: 4245

Patent and Priority Information (Country, Number, Date):

Patent: ... 20020411

Fulltext Availability:

Detailed Description

French Abstract

...40) sont transformes de maniere appropriee grace a un systeme de
transformation de contenus (10) base sur les articles d'information
des contenus en ligne et les resultats de l'analyse semantique et
conformement...

Publication Year: 2002

Detailed Description

... thedocument, andmenuinformation, thecreation
of a summary page, the creation of the lists of keywords, key
sentences etc. and links to places where the keywords etc.

appear , and the creation of the hyperlinks among the created
pages. The web contents are displayed...summary, keywords and key
sentences, the pages which contain
the lists of the keywords, key sentences etc. and the links to
the places where the keywords , key sentences etc. appear in the
document, respectively, and document fragments which are
obtained by dividing the body of...

23/3,K/20 (Item 20 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

00876811 **Image available**

SYSTEM, METHOD AND COMPUTER PROGRAM PRODUCT FOR DEVICE, OPERATING SYSTEM,
AND NETWORK TRANSPORT NEUTRAL SECURE INTERACTIVE MULTI-MEDIA MESSAGING
SYSTEME, PROCEDE ET PRODUIT PROGRAMME D'ORDINATEUR POUR APPAREIL, SYSTEME
D'EXPLOITATION ET MESSAGERIE MULTIMEDIA INTERACTIVE RESEAU, NEUTRE ET
SECURISEE

Patent Applicant/Assignee:

STORYMAIL INC, 15729 Los Gatos Boulevard, Los Gatos, CA 95032, US, US

(Residence), US (Nationality)

Inventor(s):

ILLOWSKY Daniel H, 21363 Dexter, Cupertino, CA 95014, US,
WENOCUR Michael L, 4057 Amaranta Avenue, Palo Alto, CA 94306, US,
BALDWIN Robert W, 990 Amarillo Avenue, Palo Alto, CA 94303, US,
SAXBY David B, 14946 Granite Court, Saratoga, CA 95070, US,

Legal Representative:

ANANIAN R Michael (et al) (agent), Flehr Hohbach Test Albritton & Herbert
LLP, 4 Embarcadero Center, Suite 3400, San Francisco, CA 94111-4187, US

Patent and Priority Information (Country, Number, Date):

Patent: WO 200210962 A1 20020207 (WO 0210962)

Application: WO 2001US23713 20010727 (PCT/WO US0123713)

Priority Application: US 2000627357 20000728; US 2000627358 20000728; US
2000627645 20000728; US 2000628205 20000728; US 2000706606 20001104; US
2000706609 20001104; US 2000706610 20001104; US 2000706611 20001104; US
2000706612 20001104; US 2000706613 20001104; US 2000706614 20001104; US
2000706615 20001104; US 2000706616 20001104; US 2000706617 20001104; US
2000706621 20001104; US 2000706661 20001104; US 2000706664 20001104; US
2001271455 20010225; US 2001912715 20010725; US 2001912936 20010725; US
2001912905 20010725; US 2001912773 20010725; US 2001912885 20010725; US
2001912860 20010725; US 2001912941 20010725; US 2001912901 20010725; US
2001912772 20010725

Designated States:

(Protection type is "patent" unless otherwise stated - for applications
prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR
LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL
TJ TM TR TT TZ UA UG UZ VN YU ZA ZW

(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR

(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG

(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW

(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext word Count: 169299

Patent and Priority Information (Country, Number, Date):

Patent: ... 20020207

Fulltext Availability:

Detailed Description

Publication Year: 2002

Detailed Description

... Write Key = HIVIAC (MK, Server-Subject-Name)

1. S -> C: Server-Finish

Same format as **Data** message, with the contents being the 160-bit value
SHA1(Server Nonce || Client-Nonce...

...Reuse-MK record to avoid round-trip delays.

2. C -> S: Client-Finish

Same format as **Data message**, with the contents being the 160-bit
value SHA1 (Client Nonce || Server-Nonce). This is encrypted with the
Client-write key, which is derived from **master** key.

3. Both sides confirm that the Finish records have the expected contents,
and then...

23/3,K/21 (Item 21 from file: 349)

DIALOG(R)File 349:PCT FULLTEXT

(c) 2008 WIPO/Thomson. All rts. reserv.

00809371 **Image available**

COMPUTER NETWORK INFORMATION MANAGEMENT SYSTEM AND METHOD

PROCEDE ET SYSTEME DE GESTION D'INFORMATIONS DE RESEAU INFORMATIQUE

Patent Applicant/Assignee:

TRANSCOM SOFTWARE INC, 2700 Hyde Street, San Francisco, CA 94109, US, US
(Residence), GB (Nationality)

Inventor(s):

TRIGGS Howard William Thomas, Crown Place, Shore Road, Isle of Man,
British Isles, Castletown IM9 1BF, GB,

Legal Representative:

GALLENSON Mavis S (et al) (agent), Ladas & Parry, Suite 2100, 5670
Wilshire Boulevard, Los Angeles, CA 90036-5679, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200142988 A2-A3 20010614 (WO 0142988)
Application: WO 2000US41984 20001107 (PCT/WO US0041984)
Priority Application: US 99440365 19991115

Designated States:

(Protection type is "patent" unless otherwise stated - for applications
prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE DK DM DZ EE
ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT
LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM
TR TT TZ UA UG UZ VN YU ZA ZW

(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR

(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG

(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW

(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext Word Count: 6315

Patent and Priority Information (Country, Number, Date):

Patent: ... 20010614

Fulltext Availability:

Detailed Description

English Abstract

...by gathering summary data from the information provider node
indicative of event changes at the information provider node by
information collection agents extracting information from the
information provider node based on the summary data; transmitting the
extracted information to the server; storing...

Publication Year: 2001

Detailed Description

- ... ii. Create ranked by order of occurrence the most frequent word list
(MFWL)
from the words in the RWL
- iii. Find sentences in the document containing the top 3 words in the
N1FWL
- iv. Store these sentences...

23/3,K/22 (Item 22 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

00789608

INTERACTIVE PERSONAL INFORMATION SYSTEM AND METHOD

SYSTEME D'INFORMATIONS INTERACTIF PERSONNEL ET PROCEDE CORRESPONDANT

Patent Applicant/Inventor:

ZOMMERS Oleg Kharisovich, ul. Ferganskaya, 24-179, Moscow, 109444, RU, RU
(Residence), RU (Nationality)

Legal Representative:

OBSHESTVO S OGRANICHENNOI OTVETSTVENNOSTIJU GORODISSKY I PARTNER
(agent), ul. B.Spasskaya, 25-3, Moscow, 129010, RU,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200122310 A1 20010329 (WO 0122310)
Application: WO 2000RU379 20000921 (PCT/WO RU0000379)
Priority Application: RU 99119985 19990922; US 99158562 19991008; US
2000603216 20000626

Designated States:

(Protection type is "patent" unless otherwise stated - for applications
prior to 2004)

AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB GD GE GH
GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MD MG MK MN
MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT UA UG US UZ VN YU
ZW

(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE

(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW
(EA) AM AZ BY KG KZ MD RU TJ TM
Publication Language: English
Filing Language: English
Fulltext word Count: 32297

Patent and Priority Information (Country, Number, Date):

Patent: ... 20010329

Fulltext Availability:

Detailed Description

English Abstract

...to users by having a publisher, or a multilevel structure of primary and secondary publishers, collect information items into at least one database for periodic delivery of collections of information items to users as personalized information. The collections are selected based on user profiles that are refined based on collecting and analyzing subjective...

French Abstract

...d'edition, ou d'une structure d'editeurs primaire et secondaire a multiples niveaux, qui collectent des articles d'informations dans au moins une base de donnees destinee a fournir periodiquement aux utilisateurs...

Publication Year: 2001

Detailed Description

... source document is still preserved. It is important that text of synopsis could not be found by simple removing of some words and sentences from original document.

It should be completely generated by filtering algorithm on the basis of ...

23/3,K/23 (Item 23 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

00730928 **Image available**

GENERATING PERSONALIZED USER PROFILES FOR UTILIZING THE GENERATED USER PROFILES TO PERFORM ADAPTIVE INTERNET SEARCHES
PRODUCTION DE PROFILS UTILISATEURS PERSONNALISES, UTILES POUR EXECUTER DES RECHERCHES ADAPTATIVES DANS L'INTERNET

Patent Applicant/Assignee:

MIGHTIEST LOGICON UNISEARCH INC, 2442 East 26th St., Brooklyn, NY 11235, US, US (Residence), US (Nationality)

Inventor(s):

GELLER Iiya, 2442 East 26th Street, Brooklyn, NY 11235, US

Legal Representative:

ETKIN Edward, Suite 3C, 4804 Bedford Avenue, Brooklyn, NY 11235, US

Patent and Priority Information (Country, Number, Date):

Patent: WO 200043915 A1 20000727 (WO 0043915)

Application: WO 2000US1373 20000120 (PCT/WO US0001373)

Priority Application: US 99116582 19990120; US 99422286 19991021

Designated States:

(Protection type is "patent" unless otherwise stated - for applications prior to 2004)

AE AL AU BA BB BG BR CA CN CU CZ DM EE GE HU ID IL IN IS JP KP KR LC LK
LT LV MK MN MX NO NZ PL RO SG SI SK TR TT UA UZ VN ZA

(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE

(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG

(AP) GH GM KE LS MW SD SL SZ TZ UG ZW

(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext word Count: 21317

Patent and Priority Information (Country, Number, Date):

Patent: ... 20000727

Fulltext Availability:

Claims
Publication Year: 2000

Claim

... data item segment count being representative of a number of identical segments in the corresponding data segment group of said at least one data segment group, and linking each said data...

...remote computer system, in an descending order of data item segment counts starting from a data segment group having a highest data item segment count, and recording said data segment groups and corresponding data item segment counts in said data item profile; and
(qq) storing, by the remote computer system...

...sentence mark is reached before said word count reaches a predefined word limit, storing said counted words as a sentence, restarting said word count, and repeating said step (rr) starting after a last word of said stored sentence;
and
(tt) when said word count reaches said predefined word limit, storing said counted words as a sentence, restarting said word count, and repeating said step (rr) starting after a last word of said stored sentence.
1...

23/3,K/24 (Item 24 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2008 WIPO/Thomson. All rts. reserv.

00397660 **Image available**
BUILD MESSAGE COMMUNICATION SYSTEM
SYSTEME DE COMMUNICATION CONSTRUCTEUR DE MESSAGES BINAIRE
Patent Applicant/Assignee:
CASIO COMPUTER CO LTD,
Inventor(s):
HALL Tracy R,
Patent and Priority Information (Country, Number, Date):
Patent: WO 9738403 A1 19971016
Application: WO 97JP1228 19970410 (PCT/WO JP9701228)
Priority Application: US 96631760 19960410
Designated States:
(Protection type is "patent" unless otherwise stated - for applications prior to 2004)
CN JP KR DE FR GB
Publication Language: English
Fulltext Word Count: 11137

Patent and Priority Information (Country, Number, Date):
Patent: ... 19971016
Fulltext Availability:
Detailed Description
Publication Year: 1997

Detailed Description

... to form words and build a sentence, the user simply selects buttons containing words or groups of advantage that more information can be appended with fewer bits of data information and in less time.

The present invention has advantages in building messages, The ability to build meaningful but concise sentences is made possible through preprogrammed words and phrases found in each syntax Block category. The user selects from the list of choices, and presses...

15/5/1 (Item 1 from file: 8)

DIALOG(R)File 8: Ei Compendex(R)

(c) 2008 Elsevier Eng. Info. Inc. All rts. reserv.

09620314 E.I. No: EIP03477738010

Title: Mining product reputations on the web

Author: Morinaga, Satoshi; Yamanishi, Kenji; Tateishi, Kenji; Fukushima, Toshikazu

Corporate Source: NEC Corporation, Kawasaki, Kanagawa 216-8555, Japan

Conference Title: KDD - 2002 Proceedings of the Eight ACM SIGKDD International Conference on Knowledge Discovery and Data Mining

Conference Location: Edmonton, Alta, Canada Conference Date: 20020723-20020726

Sponsor: SIGKDD; ACM Special Interest Group on Knowledge Discovery and Data

E.I. Conference No.: 61746

Source: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2002. p 341-349

Publication Year: 2002

Language: English

Document Type: CA; (Conference Article) Treatment: T; (Theoretical)

Journal Announcement: 0312w1

Abstract: Knowing the reputations of your own and/or competitors' products is important for marketing and customer relationship management. It is, however, very costly to collect and analyze survey data manually. This paper presents a new framework for mining product reputations on the Internet. It automatically collects people's opinions about target products from web pages, and it uses text mining techniques to obtain the reputations of those products. On the basis of human-test samples, we generate in advance syntactic and linguistic rules to determine whether any given statement is an opinion or not, as well as whether such an opinion is positive or negative in nature. We first collect statements regarding target products using a general search engine, and then, using the rules, extract opinions from among them and attach three labels to each opinion, labels indicating the positive/negative determination, the product name itself, and a numerical value expressing the degree of system confidence that the statement is, in fact, an opinion. The labeled opinions are then input into an opinion database. The mining of reputations, i.e., the finding of statistically meaningful information included in the database, is then conducted. We specify target categories using label values (such as positive opinions of product A) and perform four types of text mining: extraction of 1) characteristic words, 2) co-occurrence words, 3) typical sentences, for individual target categories, and 4) correspondence analysis among multiple target categories. Actual marketing data is used to demonstrate the validity and effectiveness of the framework, which offers a drastic reduction in the overall cost of reputation analysis over that of conventional survey approaches and supports the discovery of knowledge from the pool of opinions on the web. 27 Refs.

Descriptors: *Data mining; World wide web; Electronic commerce; Competition; Marketing; Customer satisfaction; Syntactics

Identifiers: Product reputations; Marketing data; Opinion labeling

Classification Codes:

723.2 (Data Processing); 723.5 (Computer Applications); 911.2 (Industrial Economics); 911.4 (Marketing)

723 (Computer Software, Data Handling & Applications); 911 (Cost & Value Engineering; Industrial Economics); 912 (Industrial Engineering & Management)

72 (COMPUTERS & DATA PROCESSING); 91 (ENGINEERING MANAGEMENT)

15/5/2 (Item 1 from file: 35)

DIALOG(R)File 35: Dissertation Abs Online

(c) 2008 ProQuest Info&Learning. All rts. reserv.

02071980 ORDER NO: AADAA-IMQ99051

A description of preverb and particle usage in Innu-aimun narrative (Labrador)

Author: Bannister, Jane

Degree: M.A.

Year: 2004

Corporate Source/Institution: Memorial University of Newfoundland
(Canada) (0306)

Adviser: Philip Branigan

Source: VOLUME 43/05 of MASTERS ABSTRACTS.

PAGE 1555. 124 PAGES

Descriptors: LANGUAGE, LINGUISTICS

Descriptor Codes: 0290

ISBN: 0-612-99051-6

Sentences with multiple preverbs and/or particles are examined in this thesis. The **data** sentences were **collected** from the first 18 stories of the Labrador Innu Text Project. Chapter 1 is an introduction to Innu-aimun grammar, with sections on previous research into word ordering, especially preverb ordering. Chapter 2 describes the patterning, use and **co - occurrence** of the ten most common preverbs in the data **sentences**. Preverbs are subdivided into modal preverbs, temporal preverbs, aspectual preverbs and other preverbs. Chapter 3 discusses 28 common particles in the data. These particles are also divided into smaller groups, including complementizers, focus particles, negative particles, adverbs, temporal and aspectual particles, particles of speaker opinion and particles with changed forms. Both chapters 2 and 3 include discussion of regular patterns of ordering of preverbs or particles. Chapter 4 is an analysis of the use of the independent or conjunct orders following negative particles. Optimality Theory is used to explain Innu data, and sentences are analyzed based on Brittain (2001, 1997). A general thesis conclusion ends chapter 4.

15/5/3 (Item 2 from file: 35)
DIALOG(R)File 35:Dissertation Abs Online
(c) 2008 ProQuest Info&Learning. All rts. reserv.

02051971 ORDER NO: AADAA-I3153573

Tracking changes in language

Author: Grothendieck, John

Degree: Ph.D.

Year: 2004

Corporate Source/Institution: Rutgers The State University of New Jersey
- New Brunswick (0190)

Director: Larry Shepp

Source: VOLUME 65/11-B OF DISSERTATION ABSTRACTS INTERNATIONAL.

PAGE 5812. 95 PAGES

Descriptors: STATISTICS ; COMPUTER SCIENCE ; ARTIFICIAL INTELLIGENCE

Descriptor Codes: 0463; 0984; 0800

ISBN: 0-496-14231-3

This thesis addresses the problems of extracting useful information from changes in a data stream including natural language. The dependencies within language complicate this analysis. Monitoring all combinations of lexical items is not computationally feasible. Even statistical tests upon single word occurrences can reveal many apparent differences. Yet the individual changes often reflect comparatively few events influencing the data. To automatically ascertain the causes of changes in the data stream requires methods for finding structure within the **set** of individual changed **items**. In this work we develop several techniques that extract such structure.

One approach utilizes word associations to cluster detected changes. The changing relationships between different lexical items—for example, the difference in correlation for word occurrence indicator variables—provide a notion of dissimilarity. A clustering algorithm with these dissimilarities as input will output groups of words that exhibit the same profile of changing **co - occurrences** with other words. This isolates novel **sentence** patterns. Changes connected to some unanticipated event cluster together, thus are readily interpreted.

Changes can be further explained by attaching them to some subset of the **data** stream. Divisive **clustering** techniques make this practical; similar **data** entries largely remain together through the clustering process. Clustering recursively reduces the complexity of the problem, stratifying the full language model into more homogeneous sub-languages. Analysis continues on these smaller, more tractable subsets. Comparing global to cluster-based tests can distinguish changes in the relative frequencies of known utterance types from novel data.

Explicit conditioning isolates the data containing particular lexical

items; standard process control tests select those features that alter in frequency. This algorithm peels away portions of the data until it detects no changes within the remainder. Implicit conditioning divides the language model so as to maximize the sample probability. This utilizes all lexical items in each data entry.

Such techniques can be combined. Together they provide an analysis package suitable for applications such as maintaining quality within an automated call center. A machine can call human attention to data that exhibits unexpected behavior in time, and help determine the nature of the observed change.

15/5/4 (Item 3 from file: 35)
DIALOG(R)File 35:Dissertation Abs Online
(c) 2008 ProQuest Info&Learning. All rts. reserv.

01335255 ORDER NO: AAD94-06319
**AUTOMATIC THESAURUS DISCOVERY VIA SELECTIVE NATURAL LANGUAGE PROCESSING: A
CORPUS BASED APPROACH**

Author: GREFENSTETTE, GREGORY THOMAS
Degree: PH.D.
Year: 1993
Corporate Source/Institution: UNIVERSITY OF PITTSBURGH (0178)
Source: VOLUME 54/09-B OF DISSERTATION ABSTRACTS INTERNATIONAL.
PAGE 4775. 344 PAGES
Descriptors: COMPUTER SCIENCE; LANGUAGE, LINGUISTICS; INFORMATION
SCIENCE
Descriptor Codes: 0984; 0290; 0723

The principal problem with information management today is organizing the ever-widening body of electronically available text. Manual techniques for filtering and structuring such information are useful for sifting through a collection of texts, but manual approaches cannot keep pace with the quantity and variety of text generated. Outside of well-funded fields such as law and medicine, there is little availability of any techniques other than simple word and stem matching for wading through this information. Such string matching techniques are thwarted, however, by the language variability problem, in which a similar idea is expressed by a variety of different words.

We defend the thesis that selective Natural Language Processing, applying subsets of known language processing techniques, over a collection of texts provides enough information to create equivalence classes between different terms, thus easing the problem of language variability. We present a method using partial syntactic analysis that allows creation of equivalence classes over any body of text and we show that the classes created by this method are more like manually-created classes than those created by document co-occurrence, sentence co-occurrence and window-based equivalence class creation techniques. Results of applying this method to information retrieval, thesaurus enrichment, and creation of automatic thesauri are also presented.

The main contributions of this thesis are the following. We describe a robust domain-independent partial parser for English which yields local syntactic contexts of words. We produce a method for using this context to create corpus-dependent similarity lists. We demonstrate that the similarities extracted by this method correspond to human similarity judgments by comparison with psychological data and by showing the overlap with manually created thesauri. We demonstrate that the overlap with manual thesauri using this syntactic context is greater than that obtained by traditional textual windowing techniques. We develop evaluation methods applicable to any corpus-based meaning extraction techniques: artificial synonyms, and gold standards measurements. We show applications of our similarity discovery techniques to information retrieval, thesaurus enrichment, and automatic thesaurus construction.

15/5/5 (Item 1 from file: 2)
DIALOG(R)File 2:INSPEC
(c) 2008 Institution of Electrical Engineers. All rts. reserv.

06300782 INSPEC Abstract Number: A9615-8760J-005, B9608-7510B-011,
C9608-7330-012
Title: Using tissue texture surrounding calcification clusters to

predict benign vs. malignant outcomes

Author(s): Thiele, D.L.; Kimme-Smith, C.; Johnson, T.D.; McCombs, M.; Bassett, L.W.

Author Affiliation: Dept. of Phys. Sci., R. Brisbane Hospital, Herston, Qld., Australia

Journal: Medical Physics vol.23, no.4 p.549-55

Publisher: AIP for American Assoc. Phys. Med.

Publication Date: April 1996 Country of Publication: USA

CODEN: MPHYA6 ISSN: 0094-2405

SICI: 0094-2405(199604)23:4L:549:UTTS;1-0

Material Identity Number: M190-96005

U.S. Copyright Clearance Center Code: 0094-2405/96/23(4)/549/7/\$10.00

Language: English Document Type: Journal Paper (JP)

Treatment: Practical (P); Experimental (X)

Abstract: The positive predictive value of mammography is between 20% and 25% for clustered microcalcifications. For very early cancers there is often a lack of concordance between mammographic signs and pathology. This study examines the usefulness of computer texture analysis to improve the accuracy of malignant diagnosis. Texture analysis of the breast tissue surrounding microcalcifications on digitally acquired images during stereotactic biopsy is used in this study to predict malignant vs. benign outcomes. 54 biopsy proven cases (36 benign, 18 malignant) are used. The texture analysis calculates statistical features from gray level co-occurrence matrices and fractal geometry for equal probability and linear quantizations of the image data. Discriminant models are generated using linear discriminant analysis and logistic discriminant analysis. Results do not differ significantly by method of quantization or discriminant analysis. Jackknife results misclassify 2 of 18 malignant cases (sensitivity 89%) and 6 of 36 benign cases (specificity 83%) for logistic discriminant analysis. From this preliminary study, texture analysis appears to show significant discriminatory power between benign and malignant tissue, which may be useful in resolving problems of discordance between pathological and mammographic findings, and may ultimately reduce the number of benign biopsies. (28 Refs)

Subfile: A B C

Descriptors: diagnostic radiography; image texture; medical image processing

Identifiers: tissue texture; calcification clusters; malignant outcome; benign outcome; digitally acquired images; stereotactic biopsy; texture analysis; statistical features; gray level co-occurrence matrices; fractal geometry; linear quantizations; image data; discriminant models; logistic discriminant analysis; mammography; medical diagnostic imaging; benign biopsies; breast tissue; computer texture analysis

Class Codes: A8760J (X-rays and particle beams (medical uses)); A8770E (Patient diagnostic methods and instrumentation); B7510B (Radiation and radioactivity applications in biomedicine); B6140C (Optical information, image and video signal processing); C7330 (Biology and medical computing); C5260B (Computer vision and image processing techniques)

Copyright 1996, IEE

15/5/6 (Item 2 from file: 2)

DIALOG(R)File 2:INSPEC

(c) 2008 Institution of Electrical Engineers. All rts. reserv.

04513958 INSPEC Abstract Number: C90001604

Title: Learning cooccurrences by using a parser

Author(s): Matsumoto, K.; Sakaki, H.; Kuroiwa, S.

Author Affiliation: KDD Kamifukuoka R&D Labs., Saitama, Japan

Conference Title: International Workshop on Parsing Technologies p. 379-88

Publisher: Carnegie Mellon Univ, Pittsburgh, PA, USA

Publication Date: 1989 Country of Publication: USA vii+467 pp.

Conference Date: 28-31 Aug. 1989 Conference Location: Pittsburgh, PA, USA

Language: English Document Type: Conference Paper (PA)

Treatment: Theoretical (T)

Abstract: Describes two methods for the acquisition and utilization of lexical cooccurrence relationships. Under these methods, cooccurrence relationships are obtained from two kinds of inputs: example sentences and the corresponding correct syntactic structure. The first of the two

methods treats a **set** of governors each **element** of which is bound to a **element** of sister nodes **set** in a syntactic structure under consideration, as a cooccurrence relationship. In the second method, a cooccurrence relationship name and affiliated attribute names are manually given in the description of augmented rewriting rules. Both methods discriminate correctness of cooccurrence by the use of the correct syntactic structure mentioned above. Experiment is made for both methods to find if thus obtained cooccurrence relationship is useful for the correct analysis. (2 Refs)

Subfile: C
Descriptors: grammars; rewriting systems
Identifiers: parser; acquisition; utilization; lexical cooccurrence;
example sentences; correct syntactic structure; cooccurrence relationship
name; affiliated attribute; rewriting rules
Class Codes: C4210 (Formal logic)

15/5/7 (Item 3 from file: 2)

DIALOG(R)File 2:INSPEC

(c) 2008 Institution of Electrical Engineers. All rts. reserv.

01912166 INSPEC Abstract Number: C76014130

Title: **Term clustering using syntactically analysed texts**

Author(s): Grishman, R.; Hirschman, L.; Sager, N.

Author Affiliation: Linguistic String Project, New York Univ., New York, NY, USA

Conference Title: Computer Science Conference /sup '/75. (Abstracts only received) p.58

Publisher: ACM, New York, NY, USA

Publication Date: 1975 Country of Publication: USA xxiv+63 pp.

Conference Sponsor: ACM

Conference Date: 18-20 Feb. 1975 Conference Location: Washington, DC, USA

Language: English Document Type: Conference Paper (PA)

Treatment: Practical (P)

Abstract: Thesauri (groupings of related words) are an important potential adjunct to information retrieval. Automatically generated thesauri have generally been based on statistical analyses of word co-occurrence within documents or sentences. Progress in mechanical syntax analysis raises the question of how information on the grammatical relation between words in a sentence could enhance thesaurus generation. The authors have developed a program to use this information, clustering nouns on the basis of the verbs with which they occur (as subject or object) and verbs on the basis of nouns (and other verbs) with which they occur. This program, applied to a small set of transformationally analyzed pharmacology texts, has yielded clusters in good agreement with the semantic word classes recognized by pharmacologists. These clusters can further be used in constructing informational formats for the text. (2 Refs)

Subfile: C

Descriptors: information retrieval; text editing; thesauri

Identifiers: syntactically analysed texts; information retrieval; thesauri; statistical analyses; pharmacology texts; term clustering

Class Codes: C7240 (Information analysis and indexing); C7250 (Information storage and retrieval)

15/5/8 (Item 1 from file: 144)

DIALOG(R)File 144:Pascal

(c) 2008 INIST/CNRS. All rts. reserv.

13780130 PASCAL No.: 98-0493369

INDEXATION DE CONNAISSANCES TEXTUELLES DANS UN SYSTEME DOCUMENTAIRE (INDEXATION OF TEXTUAL KNOWLEDGE IN AN INFORMATION RETRIEVAL SYSTEM)

VILLAIN VAN GOETHEM Marie Christine; TRIGANO Philippe, dir

Universite de Compiègne, Compiègne, Francee

Univ.: Université de Compiègne. Compiègne. FRA Degree: Th. doct.

1997-12; 1997 140 p.

Availability: INIST-T 120672; T97COMP1078 0000; RBCCN-601592101;

T97COMP1078 0000

No. of Refs.: 98 ref.

Document Type: T (Thesis) ; M (Monographic)

Country of Publication: France

Language: French Summary Language: French; English

A PARTIR D'UNE PROBLEMATIQUE QUI CONSISTAIT A DEVELOPPER UN SYSTEME DOCUMENTAIRE PERMETTANT D'INTERROGER EN LANGAGE NATUREL UNE **BASE DE TEXTES**, ON A ETE AMENE A CONCEVOIR ET A IMPLANTER UN PROTOTYPE DE SYSTEME DOCUMENTAIRE. EN ACCORD AVEC NOS HYPOTHESES DE TRAVAIL, NOTRE SYSTEME NE NECESSITE PAS DE CONNAISSANCE A PRIORI DEPENDANTE D'UN DOMAINE. NOUS AVONS TENTE DE MONTRER QU'IL EST POSSIBLE, SANS PASSER PAR UNE PHASE DE MODELISATION DES CONNAISSANCES, D'EXTRAIRE DES TEXTES UN CERTAIN NOMBRE D'INFORMATIONS UTILES DANS LE CADRE DE LA RECHERCHE D'INFORMATION. POUR CELA, NOUS AVONS PRIVILEGIE LES TECHNIQUES D'INDEXATION AUTOMATIQUE. L'ORIGINALITE DE NOTRE SYSTEME RESIDE DANS LA PRISE EN COMPTE SIMULTANEE DE DEUX ASPECTS DU DOCUMENT : - LA STRUCTURE LOGIQUE DU DOCUMENT : CE PROTOTYPE N'IMPOSE AUCUNE CONTRAINTTE PARTICULIERE SUR LA STRUCTURE DU DOCUMENT, ET PERMET DE TRAITER TOUT ENSEMBLE DE TEXTES COMPOSES DE MANIERE HIERARCHIQUE, - LES RELATIONS EXTRAITES D'UNE ANALYSE DE **COOCCURRENCES DES GROUPES NOMINAUX DU TEXTE** : UN THESAURUS EST CONSTITUE AUTOMATIQUEMENT A PARTIR DES TEXTES ANALYSES. NOTRE SYSTEME A ETE TESTE SUR DEUX CORPUS DE NATURE ASSEZ DIFFERENTE TANT PAR LEUR CONTENU QUE PAR LEUR STRUCTURE ; LES PREMIERS RESULTATS SEMBLANT ENCOURAGEANTS. LA REALISATION DE CE SYSTEME NOUS A PERMIS D'ENTREVOIR CERTAINS PROBLEMES LIES AUX TECHNIQUES DE TRAITEMENT AUTOMATIQUE DU LANGAGE NATUREL. NOUS PENSONS QUE LES TECHNIQUES STATISTIQUES ET LINGUISTIQUES SE COMBINENT AVANTAGEUSEMENT DANS LE CADRE D'UN SYSTEME DOCUMENTAIRE. CEPENDANT, ALORS QU'IL EXISTE DE NOMBREUX PROGRAMMES STATISTIQUES D'ETIQUETAGE GRAMMATICAL POUR L'ANGLAIS, PEU DE TRAVAUX ONT ETE MENES DANS CE SENS POUR LE FRANCAIS. LE SYSTEME DEVELOPPE DANS CETTE THESE FOURNIT UNE PREMIERE VERSION DU THESAURUS. NOTRE OBJECTIF FINAL EST QU'IL PUISSE ETRE CONSIDERE COMME UNE VERITABLE BASE DE CONNAISSANCE DU DOMAINE.

English Descriptors: Automatic indexing; Thesaurus; Automation; Automated processing; Document retrieval system; Information retrieval; Document structure; Linguistic analysis; Natural language; Automatic processing

French Descriptors: Indexation automatique; Thesaurus; Automatisations; Traitement automatise; Systeme documentaire; Recherche information; Structure document; Analyse linguistique; TAL; Langage naturel; Traitement automatique

Classification Codes: 001A01E01B; 205

Copyright (c) 1998 INIST-CNRS. All rights reserved.

15/5/9 (Item 2 from file: 144)

DIALOG(R)File 144:Pascal

(c) 2008 INIST/CNRS. All rts. reserv.

13012509 PASCAL No.: 97-0296001

Why words and co-words cannot map the development of the sciences

LEYDESDORFF L

Department of Science and Technology Dynamics, Nieuwe Achtergracht 166, 1018 WV Amsterdam, Netherlands

Journal: Journal of the American Society for Information Science, 1997, 48 (5) 418-427

ISSN: 0002-8231 CODEN: AISJB6 Availability: INIST-6025; 354000065515360030

No. of Refs.: 32 ref.

Document Type: P (Serial) ; A (Analytic)

Country of Publication: United States

Language: English

A restricted set of full-text articles from a sub-specialty of biochemistry was analyzed and compared in terms of co-occurrences and co-absences of words. By using the distribution of words over the sections, a clear distinction among "theoretical" "observational," and "methodological" terminology can be made in individual articles. However, at the level of the set this structure is no longer retrievable: words change both in terms of frequencies of relations with other words, and in terms of positional meaning from one text to another. These results accord with Hesse's (1980) thesis about the sciences as fluid networks. The fluidity of networks in which nodes and links may change positions is

expected to destabilize representations of developments of the sciences on the basis of co-occurrences and co-absences of words. The consequences for the lexicographical approach to generating artificial intelligence from scientific texts are discussed

English Descriptors: Scientific literature; Content analysis; Bibliometrics ; Biochemistry; **Sentence** ; Discriminant analysis; Graphics; Models; Sample; Artificial intelligence; Lexicography; **Cooccurrence** analysis; Cword; Information representation; Bibliometric map

French Descriptors: Litterature scientifique; Analyse contenu; Bibliometrie ; Biochimie; Phrase; Analyse discriminante; Representation graphique; Modele; Echantillon; Intelligence artificielle; Lexicographie; Analyse cooccurrence; Mot associe; Representation information; Carte bibliometrique

Classification Codes: 001A01A02; 205

Copyright (c) 1997 INIST-CNRS. All rights reserved.

15/5/10 (Item 1 from file: 95)
DIALOG(R)File 95:TEME-Technology & Management
(c) 2008 FIZ TECHNIK. All rts. reserv.

01026955 196095242223

Titel japanisch

(Die Verarbeitung japanischer Homonyme unter Verwendung der Information des zweiten Auftretens in einem einfachen Satz)

(Processing Japanese homonyms using information about the word **co - occurrence** in the simple **sentence**)

Takahashi, M; Yoshimura, K; Shudo, K

Fac. of Eng., Fukuoka Univ., Japan

Transactions of the Information Processing Society of Japan, v37, n6, pp998-1006, 1996

Document type: journal article Language: Japanese

Record type: Abstract

ISSN: 0387-5806

ABSTRACT:

Kana-to-kanji (phonogram-to-ideogram) conversion technology is nowadays common in Japanese word processor development. However, correct conversion without human interaction is still quite difficult because of the existence of many homonyms. We propose a new method to process homonyms on the basis of the **co - occurrence** relation between a noun and a verb in a **sentence** . Our method is based on the idea that nouns which co-occur in a simple sentence share the sentence-final verb as a governor, therefore, the most feasible candidates of kanji nouns in an input simple sentence are those each of which co-occurs with an identical verb in a simple sentence with the highest frequency. An experimental kana-to-kanji conversion using our new method for 1129 simple sentences has shown that the conversion is carried out in 93.3% of the sentences and that the accuracy is 63.0%. Our method is shown to be more effective than the ordinary method based on word occurrence frequency.

DESCRIPTORS: CHARACTER SET ; MESSAGE PROCESSING; ALGORITHM; CHARACTER RECOGNITION; CHARACTER GENERATORS
IDENTIFIERS: JAPANESE HOMONYMS; WORD **COOCCURRENCE** ; SIMPLE **SENTENCE** ; KANA TO KANJI CONVERSION; JAPANESE WORD PROCESSORS; KANJI NOUNS; HOMONYM; Textverarbeitung; japanisches Homonym; Zeichensatz

15/5/11 (Item 1 from file: 56)
DIALOG(R)File 56:Computer and Information Systems Abstracts
(c) 2008 CSA. All rts. reserv.

0000724200 IP ACCESSION NO: 200802-80-063628

Effect of dependency relationships and ordered co-occurrence of words on Japanese information retrieval (poster session)

Matsumura, Atsushi; Takasu, Atsuhiro; Adachi, Jun

National Institute of Informatics, 2-1-2, Hitotsubashi, Chiyoda-ku, Tokyo
101-8430, Japan

, p 199-200, 2000
PUBLICATION DATE: 2000

PUBLISHER: Association for Computing Machinery, Inc., One Astor Plaza, 1515
Broadway, New York, NY, 10036-5701

COUNTRY OF PUBLICATION: USA

PUBLISHER URL:

<http://portal.acm.org/citation.cfm?id=355214.355243&coll=ACM&dl=ACM&type=series&idx=SERIES9780=series&wantType=Proceedings&title=IRAL&CFID=6135546&CFTOKEN=84105396>; <http://www.acm.org/>

PUBLISHER EMAIL: SIGS@acm.org

CONFERENCE:

International workshop on Information Retrieval with Asia Languages:
Proceedings of the fifth international workshop on on Information retrieval
with Asian languages, 30 Sept.-01 Oct. 2000

DOCUMENT TYPE: Conference Paper

RECORD TYPE: Abstract

LANGUAGE: English

ISBN: 1581133006

NOTES: Hong Kong, China

DOI: 10.1145/355214.355243

FILE SEGMENT: Computer & Information Systems Abstracts

ABSTRACT:

We propose two Japanese information retrieval methods that enhance retrieval effectiveness using relationships between words. One is a method using dependency relationships between words in a **sentence**, and another is a method using the ordered **co - occurrence** information of words in a **sentence** as an approximation to the dependency relationships between them. Through retrieval experiments using the Japanese test **collection** for **information** retrieval systems NTCIR-1, we showed our two methods are superior to the TF-IDF method in retrieval effectiveness and the difference between our two methods is small. These results are independent of the **document set** and of the search topic set.

18/5/1 (Item 1 from file: 8)
DIALOG(R)File 8:EI Compendex(R)
(c) 2008 Elsevier Eng. Info. Inc. All rts. reserv.

09894160 E.I. No: EIP04248209159

Title: Information content in Medline record fields

Author: Kostoff, Ronald N.; Block, Joel A.; Stump, Jesse A.; Pfeil, Kirstin M.

Corporate Source: Office of Naval Research, Arlington, VA 22217, United States

Source: International Journal of Medical Informatics v 73 n 6 Jun 30 2004. p 515-527

Publication Year: 2004

CODEN: IJMIF4 ISSN: 1386-5056

Language: English

Document Type: JA; (Journal Article) Treatment: T; (Theoretical)

Journal Announcement: 0406w4

Abstract: Background: The authors have been conducting text mining analyses (extraction of useful information from text) of Medline records, using Abstracts as the **main data** source. For literature-based discovery, and other text mining applications as well, all records in a discipline need to be evaluated for determining prior art. Many Medline records do not contain Abstracts, but typically contain Titles and Mesh terms. Substitution of these fields for Abstracts in the non-Abstract records would restore the missing literature to some degree. Objectives: Determine how well the information content of Title and Mesh fields approximates that of Abstracts in Medline records. Approach: Select historical Medline records related to Raynaud's Phenomenon that contain Abstracts. Determine the information content in the Abstract fields through text mining. Then, determine the information content in the Title fields, the Mesh fields, and the combined Title-Mesh fields, and compare with the information content in the Abstracts. Results: Four metrics were used to compare the information content related to Raynaud's Phenomenon in the different fields: total number of phrases; number of unique phrases; content of factors from factor analyses; **content of clusters from multi-link clustering**. The Abstract field contains almost an order of magnitude more phrases than the other fields, and slightly more than an order of magnitude more unique phrases than the other fields. Each field used a factor matrix with 14 factors, and the combination of all 56 factors for the four fields represented 27 separate, but not unique, themes. These themes could be placed in two major categories, with two sub-categories per major category: Auto-immunity (antibodies, inflammation) and circulation (peripheral vessel circulation, coronary vessel circulation). All four sub-categories included representation from each field. Thus, while the focus of the representation of each field in each sub-category was moderately different, the four sub-category structure could be identified by analyzing the total factors in each field. In the cluster comparison phase of the study, the phrases used to create the clusters were the most important phrases identified for each factor. Thus, the factor matrix served as a filter for words used for clustering. While clusters were generated for all four fields, the Title hierarchy tended to be fragmented due to sparsity of the **co - occurrence** matrix that underlies the clusters. Therefore, the Title clusters were examined at only the lower levels of aggregation. The Abstract, Mesh, and Mesh + Title fields had the same first level taxonomy categories, auto-immunity and circulation. At the second level, the Abstract, Mesh, and Mesh + Title fields had the autoimmune diseases and antibodies sub-category in common. The Abstract and Mesh fields shared fascia inflammation as the other auto-immunity sub-category, while the other Mesh + Title sub-category focuses on vinyl chloride poisoning from industrial contact, and consequences of antineoplastic agents. However, in both cases, even though the words may be different, inflammation may be the common theme. Conclusions: For taxonomy generation, especially at the higher levels, each of the four fields has a similar thematic structure. At very detailed levels, the Mesh and Title fields run out of phrases relative to the Abstract field. Therefore, selection of field (s) to be employed for taxonomy generation depends on the objectives of the study, particularly the level of categorization required for the taxonomy. For information retrieval, or literature-based discovery, selection of the appropriate field again depends on the study objectives. If large queries, or large numbers of concepts or themes are desired, then the field with the largest number of technical phrases would be desirable. If queries or concepts

represented by the more accepted popular terminology is adequate, then the smaller fields may be sufficient. Because of its established and controlled vocabulary, the Mesh field lags the Title or Abstract fields in currency. Thus, the Title or Abstract fields would retrieve records with the most explicitly stated current concepts, but the Mesh field would capture a larger swath of fields that contained a concept of interest but perhaps had a wider range of specific terminology in the Abstract or Title text. In addition, this study provides the first validated estimate of the disparity in information retrieved through text mining limited to Titles and Mesh terms relative to entire Abstracts. As much of the older biomedical literature was entered into electronic databases without associated Abstracts, literature-based discovery exercises that search the older medical literature may miss a substantial proportion of relevant information. On the basis of this study, it may be estimated that up to a log order more information may be retrieved when complete Abstracts are searched. 24 Refs.

Descriptors: *Medical imaging; Information science; Data reduction; Abstracting; Antibodies; Database systems; Matrix algebra

Identifiers: Electronic databases; Data sources

Classification Codes:

461.9.1 (Immunology)

461.1 (Biomedical Engineering); 723.2 (Data Processing); 903.1 (Information Sources & Analysis); 461.9 (Biology); 723.3 (Database Systems); 921.1 (Algebra)

461 (Bioengineering); 903 (Information Science); 723 (Computer Software, Data Handling & Applications); 921 (Applied Mathematics)

46 (BIOENGINEERING); 90 (ENGINEERING, GENERAL); 72 (COMPUTERS & DATA PROCESSING); 92 (ENGINEERING MATHEMATICS)

18/5/2 (Item 2 from file: 8)
DIALOG(R)File 8: Ei Compendex(R)
(c) 2008 Elsevier Eng. Info. Inc. All rts. reserv.

08147905 E.I. No: EIP98114433514

Title: Texture classification of engineering surfaces with nanoscale roughness

Author: Grigoriev, A.Ya.; Chizhik, S.A.; Myshkin, N.K.

Corporate Source: Belarus Acad of Sciences, Gomel, Byelorussia

Source: International Journal of Machine Tools & Manufacture v 38 n 5-6 May-Jun 1998. p 719-724

Publication Year: 1998

CODEN: IMTME3 ISSN: 0890-6955

Language: English

Document Type: JA; (Journal Article) Treatment: G; (General Review)

Journal Announcement: 9812w4

Abstract: The spatial structure of the surface layer, or texture is important for surface topography characterization. In many respects a texture determines contact behavior of the rough surfaces. Despite increasing role of the precision mechanics, the texture of engineering surfaces have not been adequately investigated. In this paper pattern recognition theory is introduced to perform surface textures classification. The height-coded images obtained by atomic force microscopy were used as initial data. The images represent the surface textures of various materials formed by various processes. We take the following procedure for the texture classification. First, the texture was characterized by a matrix of co - occurrence of image contrast. Next, the matrix is transformed into feature vector by the Karhunen-Loeve transformation. The feature vector was considered as coordinates of a point in the multidimensional feature space. The location of the point depends on the peculiarities of the surface texture. The set of the points form clusters that correspond to different classes of textures. The mutual arrangement of the points and structure of the clusters were analyzed by the multidimensional scaling procedure. It was founded that there is at least four classes of surface relives. The first three of them related to the properties of surface material and the last to the process of growth and crystallization on the interface of different materials. (Author abstract) 18 Refs.

Descriptors: *Surface roughness; Pattern recognition; Textures; Mathematical transformations; Vectors

Identifiers: Engineering surfaces; Nanoscale surface roughness; Karhunen-Loeve transformation

Classification Codes:
931.2 (Physical Properties of Gases, Liquids & Solids); 723.5 (Computer Applications); 741.1 (Light/Optics); 921.3 (Mathematical Transformations); 921.1 (Algebra)
931 (Applied Physics); 723 (Computer Software); 741 (Optics & Optical Devices); 921 (Applied Mathematics)
93 (ENGINEERING PHYSICS); 72 (COMPUTERS & DATA PROCESSING); 74 (OPTICAL TECHNOLOGY); 92 (ENGINEERING MATHEMATICS)

18/5/3 (Item 1 from file: 35)
DIALOG(R)File 35:Dissertation Abs Online
(c) 2008 ProQuest Info&Learning. All rts. reserv.

01903497 ORDER NO: AADAA-I3062883
School life-histories of at-risk learning-disabled students: A retrospective study of detained and committed juveniles

Author: McCauley, Susan Diane
Degree: Ph.D.
Year: 2002
Corporate Source/Institution: The American University (0008)
Chair: Sarah Irvine-Belson
Source: VOLUME 63/08-A OF DISSERTATION ABSTRACTS INTERNATIONAL.
PAGE 2835. 184 PAGES
Descriptors: EDUCATION, SPECIAL ; SOCIOLOGY, CRIMINOLOGY AND PENOLOGY ; SOCIOLOGY, ETHNIC AND RACIAL STUDIES
Descriptor Codes: 0529; 0627; 0631
ISBN: 0-493-80587-7

Learning disabled youth represent the largest portion of special education youth in juvenile corrections. Race, gender, poverty, and urban living are all factors shown to increase the likelihood of being classified as learning disabled and identified for placement in special education, and at-risk for school failure. Research demonstrates that at-risk youth are often several years below grade level in one or more academic areas, have higher absenteeism rates, increased grade-level retention, higher dropout rates, and poorer post-school outcomes, including lower levels of meaningful employment, and higher arrest, incarceration, and recidivism rates. While research has identified a variety of factors linking academic failure and delinquency, it has been grounded in the identification of person-centered factors rather than external factors such as those found in the home, school, and community. The present study is a qualitative inquiry of adolescent learning disabled youths' perceptions of their past public schooling experiences. The study expands an examination of schooling to include the additional high-risk contexts of the youths' home and community. In-depth interviews were the primary data collection tool used for accessing the stories of these adolescent juveniles. The findings of the study suggest that youth responded and made decisions relative to their needs and socially stigmatized positioning as learning disabled students; key events co-occurring with the tasks challenges, and coping abilities of adolescence contributed to their interrupted school careers; and, persistent home, school, and community risk factors exceeded protective factors available to these adolescent students, limiting their ability to successfully adapt and respond. The distal impact of these factors on their development is demonstrated in their poor educational outcomes and increased incidences of court involvement. Students receive insufficient opportunities in their risk-prone contexts for a level of social development that lays the groundwork for adjustment and competence in adolescence and as they move toward adult roles.

18/5/4 (Item 2 from file: 35)
DIALOG(R)File 35:Dissertation Abs Online
(c) 2008 ProQuest Info&Learning. All rts. reserv.

01558442 ORDER NO: AAD13-82715
USE OF GEOGRAPHIC INFORMATION SYSTEMS (GIS) AND REMOTELY SENSED DATA TO IDENTIFY AND CHARACTERIZE THE HABITAT AND PREDICT THE DISTRIBUTION OF THE PALO DURO MOUSE, PEROMYSCUS TRUEI COMANCHE

Author: HEATON, JILL SUZANN
Degree: M.S.
Year: 1996

Corporate Source/Institution: UNIVERSITY OF NORTH TEXAS (0158)
Source: VOLUME 35/03 of MASTERS ABSTRACTS.
PAGE 749. 86 PAGES
Descriptors: BIOLOGY, ECOLOGY ; REMOTE SENSING ; PHYSICAL GEOGRAPHY
Descriptor Codes: 0329; 0799; 0368

The Palo Duro mouse, *Peromyscus truei comanche*, is known to occur only in the Texas Panhandle along the eastern "Caprock" escarpment of the Llano Estacado and is currently listed as threatened by the state of Texas. Previous studies to determine specific habitat associations that limit the species' distribution have been inconclusive. Remotely sensed data and Geographic Information Systems (GIS) were used to identify and characterize the habitat of this species, and predict the distribution of *P. t. comanche* along the Llano Estacado based on collection data, known habitat preferences, and vegetational and soil distributions. Spatial data sources employed for characterization were digitized locational information, 1:250,000 Land Use and Land Cover (LULC) data, and 1:250,000 State Soil Geographic Data Base (STATSGO) data. The co-occurrence of specific vegetation and soil types with recognized collection localities was used to identify and characterize the habitat, and predict the distribution of *P. t. comanche*.

18/5/5 (Item 1 from file: 144)
DIALOG(R)File 144:Pascal
(c) 2008 INIST/CNRS. All rts. reserv.

17677873 PASCAL No.: 06-0269589
Fuzzy predicting new association rules from current scientific literature
NAFIPS 2004 : 2004 Annual Meeting of the North American Fuzzy
Information Processing Society : Fuzzy sets in the heart of the Canadian
Rockies : Banff, Alberta, Canada, June 27-30, 2004

HUANG W; NAKAMORI Y; WANG S Y; HUYNH N V
DICK Scott, ed
School of Knowledge Science Japan Advanced Institute of Science and
Technology 1-1, Asahidai, Ishikawa 923-1292, Japan; Institute of Systems
Science Chinese Academy of Sciences, Beijing 100080, China
IEEE Systems, Man, and Cybernetics Society, United States; North American
Fuzzy Information Processing Society, United States; Institute of
Electrical and Electronics Engineers, United States
International Conference of the North American Fuzzy Information
Processing Society, 23 (Banff AB CAN) 2004
2004 450-455

Publisher: IEEE, Piscataway NJ
ISBN: 0-7803-8376-1 Availability: INIST-Y 38815; 354000138717780860
No. of Refs.: 26 ref.
Document Type: C (Conference Proceedings) ; A (Analytic)
Country of Publication: United States
Language: English

Paradoxically, the explosion of scientific information has resulted in diminishing awareness. In the face of an ever growing body of literature, disciplines are becoming increasingly specialized, while individuals and groups are becoming ever more insular. The availability of scientific bibliographies in online databases is a rich source of scientific information for scientists to support their research. In this paper, we propose a new method to predict new association rules of concepts by mining current scientific literature. In contrast to previous related research, our method's novelties are as follows: extend the antecedent and consequent of an association rule from a concept to a set of concepts; measure the relationship between two concepts not only by their co-occurrence in scientific literature, but also by their inherent relationship in knowledge bases; describe the appropriate degree of replacing a concept with its sibling; propose some indicators to distinguish various valid changes of existing association rules. The predicted new association rules can serve researchers as major repositories of candidates for new research themes, as impetus for inspiration impetus, or as hypotheses to be tested in future.

English Descriptors: Data mining; Scientific technical information;
Availability; Database; Information source; Statistical association;
Bibliography; Cooccurrence analysis; Hypothesis test; Fuzzy logic;
Association rule

French Descriptors: Fouille donnee; Information scientifique technique;
Disponibilite; **Base** donnee; Source **information** ; Association
statistique; Bibliographie; Analyse **cooccurrence** ; Test hypothese;
Logique floue; Regle association

Classification Codes: 001D02B07B; 001D02C

Copyright (c) 2006 INIST-CNRS. All rights reserved.

18/5/6 (Item 2 from file: 144)
DIALOG(R)File 144: Pascal
(c) 2008 INIST/CNRS. All rts. reserv.

16512821 PASCAL No.: 04-0158782
**Discovering exceptional information from customer inquiry by association
rule miner**

DS 2003: discovery science : Sapporo, 17-19 October 2003
SHIMAZU Keiko; MOMMA Atsuhito; FURUKAWA Koichi
GRIESER Gunter, ed; TANAKA Yuzuru, ed; YAMAMOTO Akihiro, ed
Information Media Laboratory, Corporate Research Group, Fuji Xerox Co.,
Ltd, 430 Sakai Nakai-machi Ashigarakami-gun Kanagawa 259-0157, Japan;
Graduate School of Media and Governance, Keio University, 5322 Endo
Fujisawa-shi Kanagawa 252-8520, Japan
Discovery science. International conference, 6 (Sapporo JPN) 2003-10-17
Journal: Lecture notes in computer science, 2003, 2843 269-282
ISBN: 3-540-20293-5 ISSN: 0302-9743 Availability: INIST-16343;
354000117768960210

No. of Refs.: 23 ref.
Document Type: P (Serial); C (Conference Proceedings) ; A (Analytic)
Country of Publication: Germany
Language: English

This paper reports the results of our experimental study on a new method of applying an association rule miner to discover useful information from a text database. It has been claimed that association rule mining is not suited for text mining. To overcome this problem, we propose (1) to generate a sequential **data set** of words with dependency structure from a Japanese text database, and (2) to employ a new method for extracting meaningful association rules by applying a new rule selection criterion. Each inquiry was converted to a list of word pairs, having dependency relationship in the original sentence. The association rules were acquired regarding each pair of words as an item. The rule selection criterion derived from our principle of giving heavier weights to **co - occurrence** of multiple items than to single item occurrence. We regarded a rule as important if the existence of the items in the rule body significantly affected the occurrence of the item in the rule head. Based on this method, we conducted experiments on a customer inquiry database in a call center of a company and successfully acquired practical meaningful rules, which were not too general nor appeared only rarely. Also, they were not acquired by only simple keyword retrieval. Additionally, inquiries with multiple aspects were properly classified into corresponding multiple categories. Furthermore, we compared (i) rules obtained from a sequential **data set** of words with dependency structure, which we propose in this paper, and those without dependency structure, as well as (ii) rules acquired through the association rule selection criterion and those through the conventional criteria. As a result, discovery of meaningful rules increased 14.3-fold in the first comparison, and we confirmed that our criterion enables to obtain rules according to the objectives more precisely in the second comparison.

English Descriptors: Artificial intelligence; Data mining; Text; Database;
Useful information; Knowledge discovery; Sequential; Japanese; Selection
criterion; **Cooccurrence** analysis; Selection rule

French Descriptors: Intelligence artificielle; Fouille donnee; Texte; **Base**
donnee; **Information** utile; Decouverte connaissance; Sequentiel;
Japonais; Critere selection; Analyse **cooccurrence** ; Regle selection

Classification Codes: 001D02C04

Copyright (c) 2004 INIST-CNRS. All rights reserved.

18/5/7 (Item 3 from file: 144)
DIALOG(R)File 144:Pascal
(c) 2008 INIST/CNRS. All rts. reserv.

14395481 PASCAL No.: 00-0049798
The epiphyte vegetation of *Annona glabra* on Barro Colorado Island, Panama
ZOTZ G; BERMEJO P; DIETZ H
Lehrstuhl fuer Botanik II der Universitaet wuerzburg,
Julius-von-Sachs-Platz 3, 97082 wuerzburg, Germany; Smithsonian Tropical
Research Institute, Apdo. 2072, Balboa, Panama
Journal: Journal of biogeography, 1999, 26 (4) 761-776
ISSN: 0305-0270 CODEN: JBIODN Availability: INIST-15698;
354000080954330050

No. of Refs.: 1 p.1/4
Document Type: P (Serial) ; A (Analytic)
Country of Publication: United Kingdom
Language: English

Aim Information on the community composition, structure, and dynamics of epiphyte vegetation is scarce. A survey of the epiphytes occurring on all individuals of one particular host tree species in a well-studied neotropical research site allowed us a comparison of the epiphyte flora of this tree with the local epiphyte flora, the analysis of spatial distribution patterns and the use of these patterns as indications for changes in time. In the future, our results can be used as a **baseline data** - set for the direct observation of the long-term dynamics in epiphyte communities. Location The study was conducted on Barro Colorado Island (BCI), Panama. Methods we recorded all individuals of the vascular epiphytes growing on *Annona glabra* L., a flood-tolerant, multiple-stemmed tree, which is restricted to the shoreline of BCI. Data on tree biometrics, epiphyte species, and epiphyte abundances were collected for more than 1200 trees. Results In total, we encountered almost 15,000 epiphytic individuals in sixty-eight species, corresponding to more than one third of the entire epiphyte flora of Barro Colorado Island. The component species differed strongly in abundance: the four most important species accounted for >75% of all individuals. In most cases, the same four species were also the first to colonize a tree (=phorophyte). Colonization patterns indicated no replacement of early colonizers by late arrivals. Species richness and epiphyte abundances showed a positive correlation with the size and the density of the host trees. All species showed a highly clumped distribution and the physiognomy of epiphyte communities of individual trees was dominated either by one or several of the four most common species or by a set of frequently co - occurring tank bromeliads. Other species were dominant only in exceptional cases. Most species were always rare. A distance effect on community composition was mostly confined to a local scale with an increased similarity in the species assemblage of stems of a tree v. neighbouring trees. Main conclusions The epiphytes on a single small phorophyte species may encompass a surprisingly large proportion of the local epiphyte flora. The observations that most tree crowns are inhabited by a single or only very few species, and that all epiphyte species show highly clumped distributions suggest a predominance of very local dispersal within a tree crown, which is only infrequently interrupteded by successful long-distance dispersal between crowns.

English Descriptors: Species diversity; Population density; Species richness; Floral survey; Spatial distribution; Vegetation; Epiphyte; Pruning; Distance; Tree; Island; Panama
Broad Descriptors: Central America; America; Annonaceae; Dicotyledones; Angiospermae; Spermatophyta; Community structure; Amerique Centrale; Amerique; Annonaceae; Dicotyledones; Angiospermae; Spermatophyta; Structure communautaire; America central; America; Annonaceae; Dicotyledones ; Angiospermae; Spermatophyta

French Descriptors: Diversite especes; Densite population; Richesse specifique; Inventaire floristique; Repartition spatiale; Vegetation; Epiphyte; Taille plante; Distance; Arbre; Ile; Panama; *Annona glabra*

Classification Codes: 002A14B04B
Copyright (c) 2000 INIST-CNRS. All rights reserved.

18/5/8 (Item 4 from file: 144)
DIALOG(R)File 144:Pascal
(c) 2008 INIST/CNRS. All rts. reserv.

13465853 PASCAL No.: 98-0162486
Longitudinal study of co - occurring psychiatric disorders and substance use

BROOK J S; COHEN P; BROOK D W
Department of Community Medicine, Mount Sinai School of Medicine, New York, United States; Department of Psychiatry, Columbia University, New York, United States

Journal: Journal of the American Academy of Child and Adolescent Psychiatry, 1998, 37 (3) 322-330

ISSN: 0890-8567 CODEN: JAAPEE Availability: INIST-2261;
354000078370820160

No. of Refs.: 34 ref.

Document Type: P (Serial) ; A (Analytic)

Country of Publication: United States

Language: English

Objective: To examine temporal priority in the relationship between psychiatric disorders and drug use. Method: Psychiatric assessments and drug use were completed at three different points in time, spanning 9 years. Structured interviews were administered to a cohort of youths and their mothers. Subjects were selected on the basis of their residence in either of two counties in upstate New York. The sample was predominantly white male and female youths, aged 1 through 10 years upon initial collection of data. Psychiatric diagnoses were assessed by a supplemented version of the Diagnostic Interview Schedule for Children version 1, using computer algorithms designed to match DSM-III-R criteria to combine information from mothers and youths. Substance use information was obtained in the interviews. Results: A significant relationship was found to exist between earlier adolescent drug use and later depressive and disruptive disorders in young adulthood, controlling for earlier psychiatric disorders. Earlier psychiatric disorders did not predict changes in young adult drug use. Conclusions: Implications for policy, prevention, and treatment include (1) more medical attention needs to be given to the use of legal and illegal drugs; and (2) a decrease in drug use may result in a decrease in the incidence of later psychiatric disorders.

English Descriptors: Concomitant disease; Mental disorder; Alcoholism; Drug addiction; Tobacco smoking; Follow up study; Infant; Preschool age; School age; Child; Preadolescent; Adolescent; Young adult

Broad Descriptors: Human; Homme; Hombre

French Descriptors: Association morbide; Trouble psychiatrique; Alcoolisme; Toxicomanie; Tabagisme; Etude longitudinale; Nourrisson; Age prescolaire; Age scolaire; Enfant; Preadolescent; Adolescent; Adulte jeune

Classification Codes: 002B18C05D

Copyright (c) 1998 INIST-CNRS. All rights reserved.

18/5/9 (Item 5 from file: 144)
DIALOG(R)File 144:Pascal
(c) 2008 INIST/CNRS. All rts. reserv.

10355187 PASCAL No.: 92-0558647

Autism and tuberous sclerosis

SMALLEY S L; TANGUAY P E; SMITH M; GUTIERREZ G

UCLA, dep. psychiatry, Los Angeles CA, USA

Journal: Journal of autism and developmental disorders, 1992, 22 (3)
339-355

ISSN: 0162-3257 CODEN: JADDDQ Availability: INIST-15018;
354000030229030020

No. of Refs.: 1 p.1/2

Document Type: P (Serial) ; A (Analytic)

Country of Publication: USA

Language: English

Autism is a behavior disorder with genetic influences indicated from twin and family studies and from the cooccurrence of autism with known genetic

disorders. Tuberous sclerosis complex (TSC) is a known genetic disorder with behavioral manifestations including autism. A literature review of these two disorders substantiates a significant association of autism and TSC with 17-58% of TSC subjects manifesting autism and 0.4-3% of autistic subjects having TSC. In initial data collected on 13 TSC probands and 14 autistic probands in our family study of autism and TSC, we identified 7 TSC subjects with autism

English Descriptors: Autism; Psychosis; Developmental disorder; Concomitant disease; Review; Bourneville syndrome; Tumor; Nervous system diseases; Family study; Inheritance; Child
Broad Descriptors: Human; Homme; Hombre

French Descriptors: Autisme; Psychose; Trouble developpement; Association morbide; Article synthese; Phacomatose Bourneville; Tumeur; Systeme nerveux pathologie; Etude familiale; Determinisme genetique; Enfant

Classification Codes: 002B18D04B

18/5/10 (Item 1 from file: 266)

DIALOG(R)File 266:FEDRIP

Comp & dist by NTIS, Intl Copyright All Rights Res. All rts. reserv.

00653855

IDENTIFYING NO.: 1K01MH079945-01 AGENCY CODE: CRISP

A Cognitive-Behavioral Intervention for Children with Autism Spectrum Disorders

PRINCIPAL INVESTIGATOR: WILLIAMS, SUSAN K

ADDRESS: SWILLIAMS25@VCU.EDU VIRGINIA TREAT CTR FOR CHILDREN 515 NORTH 19TH STREET RICHMOND, VA 23298

PERFORMING ORG.: VIRGINIA COMMONWEALTH UNIVERSITY, RICHMOND, VIRGINIA

SPONSORING ORG.: NATIONAL INSTITUTE OF MENTAL HEALTH

DATES: 2009/01/07 TO 2007/31/12 FY : 2007 TYPE OF AWARD: New Award (Type 1)

SUMMARY: DESCRIPTION (provided by candidate): I am applying for the Mentored Research Scientist Development Award (K01) program to promote my growth as an independent scientist. My goal with this award is to obtain the advanced training I need to develop an evidence-based treatment manual for a social skills intervention for children with autism spectrum disorders (ASD), and subsequently apply for an R01 to test these interventions in clinical trials. Children and adolescents with ASD who are cognitively 'high-functioning' frequently present with symptoms of anxiety, in addition to the core impairment in social interaction. This anxiety can interfere with their ability to integrate into mainstream academic environments, undermine their use of appropriate social skills in natural peer contexts, and impede their overall development. Thus, anxiety can be seen as compounding the social disability inherent in spectrum disorders, preventing otherwise able children from reaping the maximum benefit from interventions targeting social skill development. The K01 career development aims will allow me to gain additional instruction, mentoring and experience in: (1) assessment and treatment of ASD and childhood anxiety; (2) the design of psychosocial treatment manuals; (3) methods and statistical techniques appropriate for the design and conduct of randomized controlled trials of psychosocial interventions; and (4) responsible and ethical conduct of research. These training objectives relate directly to my research plan, the ultimate goal of which is to develop an evidence-based, efficacious treatment program for children with ASD that targets social skill development and anxiety reduction. The aims of this research are: (1) to develop an alpha version of a treatment manual that addresses social skill development and co-occurring anxiety in school-age children and adolescents with ASD; (2) to pilot strategies comprising the treatment manual with a small group (n=5) of children to refine intervention strategies and delivery; (3) to collect preliminary data on the short-term efficacy, as well as feasibility, of this structured manual-based treatment in a sample (n=24) of children with ASD complicated by anxiety; and (4) to develop a grant application to conduct a larger scale efficacy study of the treatment curriculum. Through this training and research plan, I will be well-positioned to carry out independent investigations designed to translate an empirical understanding of anxiety and social disability in children with high-functioning ASD into novel treatment approaches.

DESCRIPTORS: child psychology; adolescence (12-20); middle childhood (6-11); clinical trial; human subject; anxiety; autism; social behavior disorder; social behavior; handbook; human therapy evaluation; cognitive behavior therapy; clinical research; behavioral /social science research tag; Asperger syndrome; therapy design /development

18/5/11 (Item 2 from file: 266)

DIALOG(R)File 266:FEDRIP

Comp & dist by NTIS, Intl Copyright All Rights Res. All rts. reserv.

00653046

IDENTIFYING NO.: 1R03MH075833-01A2 AGENCY CODE: CRISP

Toward a Rat Model of Alcohol Abuse in Schizophrenia

PRINCIPAL INVESTIGATOR: CHAU, DAVID

ADDRESS: david.t.chau@dartmouth.edu David Thanh Chau Dartmouth Medical School Lebanon, NH 03756

PERFORMING ORG.: DARTMOUTH COLLEGE, HANOVER, NEW HAMPSHIRE

SPONSORING ORG.: NATIONAL INSTITUTE OF MENTAL HEALTH

DATES: 2008/15/07 TO 2004/30/09 FY : 2007 TYPE OF AWARD: New Award (Type 1)

SUMMARY: DESCRIPTION (provided by applicant): Alcohol use disorder commonly occurs among patients with schizophrenia and contributes greatly to the morbidity of schizophrenia. Patients with schizophrenia tend to consume modest quantities of alcohol on a regular basis and are less likely to develop alcohol dependence than alcohol abuse, but even this modest use of alcohol dramatically worsens their symptoms and decreases their overall functioning. Green (co- investigator) and colleagues have suggested that such moderate alcohol use may, however, transiently ameliorate a brain reward circuit deficiency that underlies alcohol use disorder in these patients. Unfortunately, available treatments for co - occurring alcohol use disorder in schizophrenia are very limited. This revised R03 proposal seeks to begin a line of research toward the development of such an animal model of alcohol use disorder in schizophrenia, an animal that exhibits characteristics of schizophrenia, and like patients with schizophrenia, also drinks at least moderate amounts of alcohol. To develop this animal model, we propose to use, as a base, a rat with a neonatal ventral hippocampal lesion (the NVHL rat), a well-established animal model of schizophrenia, a rodent that as an adult exhibits requisite characteristics of schizophrenia, demonstrates abnormalities in its brain reward circuit, and, interestingly, has recently been shown to exhibit increased cocaine self-administration. Our preliminary data in a small group of adult NVHL rats also suggest that this rat will voluntarily drink at least moderate amounts of alcohol. This revised research proposal seeks to further probe the potential role of the NVHL rat as an animal model of schizophrenia and comorbid alcohol use disorder. Using free-access conditions, we will: (1) compare the amount and preference of alcohol drinking [and blood alcohol level] in NVHL rats versus sham-operated rats; and (2) compare the size, frequency and temporal distribution of alcohol drinking bouts in NVHL rats versus sham-operated rats. If NVHL rats can be differentiated from sham rats according to these measures, we plan to continue research with NVHL rats in subsequent studies to: (1) explore mechanisms mediating alcohol drinking in these animals (e.g., to address the question of whether alcohol use serves to transiently ameliorate a deficit in brain reward functioning); and (2) screen medications that might be able to decrease alcohol drinking in this rat. Ultimately, we expect to translate the findings from our studies with the NVHL rat into studies involving human subjects, with the long- term goal of this research to find novel medications to treat patients with schizophrenia and alcohol use disorder, and thus to improve the outcome of these patients. Alcohol use disorder occurs commonly among patients with schizophrenia and greatly worsens the overall functioning of these patients. This research seeks to develop an animal model of alcohol use disorder in schizophrenia, an animal model that exhibits schizophrenia-like characteristics as well as increased alcohol drinking. This animal model, when developed, will be used: (1) to elucidate the underlying basis of alcohol use disorder in patients with schizophrenia; and, (2) to develop novel medications to limit alcohol use in these patients.

DESCRIPTORS: alcoholism /alcohol abuse; gamma aminobutyrate; laboratory rat; hippocampus; experimental brain lesion; self medication; disease /disorder model; dopamine; behavior test; preference; reinforcer; antipsychotic agent; schizophrenia; comorbidity; behavioral /social science

research tag; substance abuse related behavior

18/5/12 (Item 3 from file: 266)

DIALOG(R)File 266:FEDRIP

Comp & dist by NTIS, Intl Copyright All Rights Res. All rts. reserv.

00613937

IDENTIFYING NO.: 1R01DA023072-01A1 AGENCY CODE: CRISP

Development of a Treatment Adherence Program for Bipolar Substance Abusers

PRINCIPAL INVESTIGATOR: MILLER, IVAN W.

ADDRESS: Ivan.Miller@brown.edu Butler Hospital Providence, RI 02906

PERFORMING ORG.: BUTLER HOSPITAL (PROVIDENCE, RI), PROVIDENCE, RHODE ISLAND

SPONSORING ORG.: NATIONAL INSTITUTE ON DRUG ABUSE

DATES: 2009/30/07 TO 2007/31/10 FY : 2007 TYPE OF AWARD: New Award (Type 1)

SUMMARY: DESCRIPTION (provided by applicant): There is a substantial co - occurrence between substance use disorders and bipolar disorder. Bipolar disorder is five to eight times more likely to occur in patients with substance use disorders than in the general population. Conversely, rates of substance use disorders in bipolar samples have been reported to be as high as 60% or more. This comorbidity is associated with an earlier age of illness onset, increased symptom severity, greater tendency for violence, higher rates of psychiatric hospitalization, slower time to remission of acute mood episodes, poorer response to lithium treatment, and increased suicidality and mortality rates. Although treatment nonadherence is a significant problem in both substance use and bipolar disorders independently, the co - occurrence of these conditions is related to even poorer compliance rates. Further, research indicates that bipolar substance abusers have a worse course of illness compared to noncomorbid patients, and nonadherence is the most consistent predictor of these poor outcomes. To date, there is very little research on behavioral interventions specifically designed to improve treatment adherence in this high-risk, comorbid population. The present proposal is designed to meet the objectives of Stage I of the NIDA Behavioral and Integrative Therapies Development Program (PA-07-111). Stage I is the initial stage of treatment development research: to formulate new behavioral therapies; to operationally define therapy manuals and procedures; and to pilot test and refine new therapies. The overall aim of this study is to develop the Integrated Treatment Adherence Program (ITAP), which is designed as an adjunctive intervention for improving treatment adherence (broadly defined) in bipolar substance abusers, and to collect preliminary data on the feasibility, acceptability, and initial efficacy of the program. More specifically, we propose the following major aims: 1. To develop a comprehensive treatment manual for ITAP - an innovative, multi-modal intervention that combines motivational, family, and telephone-based strategies - by conducting a small open trial (n = 15) with patients with drug dependence and bipolar disorder. 2. To conduct a randomized controlled pilot study in a sample (n = 60) of patients initially hospitalized with comorbid substance dependence and bipolar disorder by comparing ITAP to treatment as usual to estimate relevant treatment parameters (e.g., acceptability, preliminary efficacy), including HIV risk behaviors. This pilot study will lay the groundwork for a larger clinical trial (Stage II) evaluating the efficacy of this new treatment for improving adherence in bipolar substance abusers.

18/5/13 (Item 4 from file: 266)

DIALOG(R)File 266:FEDRIP

Comp & dist by NTIS, Intl Copyright All Rights Res. All rts. reserv.

00557483

IDENTIFYING NO.: 0187777 AGENCY CODE: AGRIC

HERBIVORE-MEDIATED INDIRECT EFFECTS OF AN EXOTIC THISTLE ON NATIVE THISTLES.

spatial distribution

ASSOCIATE INVESTIGATORS: Louda, S. M.

PERFORMING ORG.: UNIVERSITY OF NEBRASKA, SCHOOL OF BIOLOGICAL SCIENCES, LINCOLN, NEBRASKA 68583

TYPE OF AWARD: NRI COMPETITIVE GRANT |c c

SUMMARY: The overall objective of this project is evaluate the interaction between an invasive exotic weed and two related native plant species mediated by a shared insect herbivore, a deliberately released biocontrol weevil, *Rhinocyllus conicus*, in prairie rangelands. The weevil is here and having significant impacts on prairie species. Can this impact be managed and, if so, how? The preliminary data suggested that *R. conicus* reduced its use and impact on native species in the vicinity of its preferred, exotic host plant, musk thistle (*Carduus nutans* spp.). Thus, the objectives of the research are to: 1) evaluate the generality of this observation in Nebraska grasslands, 2) develop a

better mechanistic, experimentally-based understanding of insect-mediated indirect interactions between plant species within this system, and 3) examine the applicability of that new data to managing the impact of *R. conicus* on sparse or rare native plant species. The study is designed to answer three fundamental questions: 1) how are seed losses of native thistle species related to ecological circumstances, such as proximity to stands of the targeted weed, musk thistle, and surrounding vegetation; 2) are plant co - occurrence and observed levels of impact causally related; and, 3) can the ecological factors be manipulated to minimize negative impacts on native

species. The aim is to improve our basic understanding of herbivore-mediated indirect interactions between co - occurring plants and apply that understanding to science-based management of non-target effects associated with the biological control of invasive plant species, such as thistles. The study entails both data collection on the pattern of injury inflicted by *Rhinocyllus conicus* on native thistles in prairie grasslands and the response of *R. conicus* to native species in experimentally planted arrays. The hypotheses to be tested are that co - occurrence of the native species with musk thistle: (H1) has no effect on seed loss of the native, or (H2) decreases seed loss

by the native ("associational defense"), or (H3) increases seed loss by the native ("associational susceptibility"). The patterns will be documented in relation to proximity to musk thistle (*Carduus nutans* ssp. *leioophyllus*) stands as well as variation in weevil densities and identity of the ambient plant community. The experiments will determine directly the degree to which native plant use is influenced by proximity to stands of the targeted, preferred weed species. PR factors, specifically local density and proximity of the exotic weed musk thistle (*Carduus nutans*), affect non-target damage to native species such as wavyleaf thistle (*Cirsium undulatum*), by *Rhinocyllus conicus*, a

European weevil used as a biocontrol agent for musk thistle. Theory predicts that co - occurrence with musk thistle could increase or decrease *R. conicus* damage to native thistles. Damage may be less where wavyleaf and musk thistles co-occur (associational defense), if weevils are drawn from the native by their preferred host, musk thistle. Alternately, co - occurrence may increase damage to native thistles (associational susceptibility), if weevils on musk thistle spillover onto natives. Our results should help reduce *R. conicus* non-target effects by providing data needed to evaluate two management strategies: 1) reduce musk thistle abundance to minimize weevil spillover onto

natives vs. 2) use musk thistles as a 'trap crop' to draw weevils from natives. We quantified *R. conicus* use of the wavyleaf thistle in three ways: regional surveys (2001-2003), experimental manipulations of *R. conicus* (2002-2003), and quantification of *R. conicus* oviposition in relation wavyleaf density on loess soils (2004) in southwest Nebraska. Earlier data from the Sand Hills, where musk thistle is very rare, showed more *R. conicus* damage to wavyleaf thistles in dense patches (>5 stems in a 3 m radius) than to isolated plants (>20 m from any bolting thistle). Our results pr

PROGRESS REPORT SUMMARY: Russell, F. L. and S. M. Louda. 2004. Phenological synchrony affects interaction strength of an exotic weevil with Platte thistle, a native host plant. *Oecologia* 139:525-534. Rand, T. A., F. L. Russell and S. M. Louda. 2004. Local vs. landscape scale indirect effects of an invasive weed on native plants. *Weed Technology* 18:1250-1254. Louda, S. M., T. A. Rand, F. L. Russell and A. E. Arnett. 2005.

Assessment of Ecological Risks in Biocontrol: Input from Retrospective Analyses. Biological Control: in press. Russell, F. L. and S. M. Louda. 2005. Insect abundance, phenology and associational defense influence floral herbivory by an invasive insect. *Oecologia*. In review. Russell, F. L., S. M. Louda and T. A. Rand. 2005. Variation in herbivore-mediated indirect effects of an invasive plant on a native plant. In draft.

Russell, F. L. and S. M. Louda. 2006. Spatial variation in *Rhinocyllus conicus* response to density of an adopted native host plant, *Cirsium undulatum*. In prep. Russell, F. L. and S. M. Louda. 2006. Does weed density explain variation in *Rhinocyllus conicus* damage to a target host plant, *Carduus nutans* (musk thistle)? In prep.

DESCRIPTORS: herbivores; insects; environmental impact; defense mechanisms; invasive species; plant competition; weed control; non target organisms; biological control (weeds); prairies; rangelands; *Rhinocyllus conicus*; *Cirsium*; *Carduus nutans*; risk management; plant ecology; insect ecology; exotic plants; native plants; ecosystem management; plant damage; spatial distribution

18/5/14 (Item 5 from file: 266)

DIALOG(R) File 266: FEDRIP

Comp & dist by NTIS, Intl Copyright All Rights Res. All rts. reserv.

00507665

IDENTIFYING NO.: 179941; 0001; 648 AGENCY CODE: VA

Atypical Antipsychotic Use and Smoking Cessation in Those with Bipolar Disorder and Schizophrenia

PRINCIPAL INVESTIGATOR: Matthews, Annette M, M.D.

PERFORMING ORG.: Department of Veterans Affairs, Medical Center, Portland, OR

SPONSORING ORG.: Department of Veterans Affairs, Research and Development (15), 810 Vermont Ave. N.W., Washington, D.C. 20420 United States of America

DATES: 20070101

SUMMARY: SMOKING; ANTIPSYCHOTIC AGENTS; BIPOLAR DISORDER; SCHIZOPHRENIA

OBJECTIVES: To determine if atypical antipsychotic use is associated with a decreased rate of smoking cessation in those with bipolar disorder as compared to those with schizophrenia.

PLAN: We will collect data on patients treated for bipolar disorder in the Veterans Administration VISN 20 between January 2000 and December 2005. VISN 20 includes 8 medical centers and 17 community-based outpatient clinics (CBOCs) distributed throughout Alaska, Washington, Oregon, and Idaho. We require the data by downloading it from the VISN 20 data warehouse into a local database using structured query language (SQL) queries to organize it and export to SPSS 14.0 for analysis. We will compare those with schizophrenia on atypical antipsychotics with those with bipolar disorder on atypical antipsychotics who smoke.

METHODS: We will use the Cox proportional hazards models to compare the two groups. We will examine time from baseline, January 2000, until smoking cessation, where smoking cessation is measured by the subjects answer an annual required smoking cessation alert. We will control for participant characteristics such as age, medications, and co-occurring conditions. Diagnosis (bipolar disorder versus schizophrenia) will be included as a predictor. Medications will be entered into the models as time-varying co-variates. From a preliminary data analysis we know that we will be able to look at data for at least 300 people in each group of interest. Portland VA Medical Center Institutional Review Board approval for data collection and analysis will be obtained before the data is collected.

FINDINGS TO DATE: none.

*** PDS Report: Initial; Report Date: 01/01/07; Submitted: 06/21/07 ***

Initial Report

DESCRIPTORS: SMOKING; ANTIPSYCHOTIC AGENTS; BIPOLAR DISORDER; SCHIZOPHRENIA

18/5/15 (Item 6 from file: 266)

DIALOG(R) File 266: FEDRIP

Comp & dist by NTIS, Intl Copyright All Rights Res. All rts. reserv.

00506271

IDENTIFYING NO.: 182706; 0052; 640 AGENCY CODE: VA

Archival Data Analysis of Traumatic Brain Injury and Co-existing Psychiatric Illness in Veterans

PRINCIPAL INVESTIGATOR: Yesavage, Jerome A., M.D.

PERFORMING ORG.: Department of Veterans Affairs, Medical Center, Palo Alto, CA

SPONSORING ORG.: Department of Veterans Affairs, Research and Development

(15), 810 Vermont Ave. N.W., Washington, D.C. 20420 United States of America

DATES: 20071114

SUMMARY: MENTAL DISORDERS; DATABASE; HEAD INJURY; VETERANS

The purpose of this archival data analysis and chart review study is to examine the relationship between traumatic brain injury (TBI) and co-existing psychiatric illness in veterans in the VA Local Vista database and/or the National Patient Care Database (NPCD). The goal is to characterize the co-existing psychiatric illness in this population, to improve the planning and development of targeted programs for early detection and intervention of co-existing psychiatric illness in this patient population. The aims of the proposed data analysis study are to: (1) evaluate the sociodemographic characteristics and VA health service utilization of veterans who have sustained a documented TBI as reflected by VA medical records; (2) determine the prevalence rate of co-occurring psychiatric diagnoses in veterans with TBI increase as a function by year.

RESEARCH PLAN AND METHODS

The proposed study will be a prospective cohort design and an archival data analysis. In the proposed study, we will analyze data from the VA Local Vista database and/or the National Patient Care Database (NPCD). Participants between the ages of 18 and 89 years, of any race or ethnicity, and who meet our inclusion and exclusion criteria, will be included in the study. In addition, we will include demographically matched veteran control participants. Primary data to be collected are brain injury-related and psychiatric diagnostic categories. Other primary data will include TBI clinical reminder questions. Descriptive statistics will be used to analyze sociodemographic variables, including gender, ethnicity, age, marital status, armed forces component, and health service utilization. Additional statistical analyses will be used as needed to compute secondary analyses (i.e., ANOVA, multiple regression).

CLINICAL RELEVANCE

Traumatic brain injury (TBI) is reported to be the most common consequence of combat-related injuries among surviving U.S. soldiers in the Operation Enduring Freedom (OEF) and Operation Iraqi Freedom (OIF) conflicts. The increasing number of TBI trauma survivors is a high priority, high cost area for the VA. Research also suggests that TBI may cause decade-long or even permanent vulnerability to psychiatric illness in some individuals. The findings are critical to development of targeted programs for early identification and intervention for veterans who may have TBI and/or co-existing psychiatric illness. As the veteran population ages, disability as a consequence of TBI will become a significant health care issue in the coming decades in the treatment of older veterans (e.g., Vietnam, Korea, WWI era) in the VA. Given that the majority of TBI have been blast-related, the generalizability beyond the VA is questionable.

*** PDS Report: Initial; Report Date: 11/14/07; Submitted: 11/19/07 ***
Initial Report

DESCRIPTORS: MENTAL DISORDERS; DATABASE; HEAD INJURY; VETERANS

18/5/16 (Item 7 from file: 266)

DIALOG(R) File 266: FEDRIP

Comp & dist by NTIS, Intl Copyright All Rights Res. All rts. reserv.

00501674

IDENTIFYING NO.: 152055; 0001; 605 AGENCY CODE: VA

Unified Psychogeriatric Biopsychosocial Evaluation and Treatment (UPBEAT)

PRINCIPAL INVESTIGATOR: Blow, Frederic C., Ph.D.

PERFORMING ORG.: Department of Veterans Affairs, Medical Center, Loma Linda, CA

SPONSORING ORG.: Department of Veterans Affairs, Research and Development (15), 810 Vermont Ave. N.W., Washington, D.C. 20420 United States of America

DATES: 19950215

SUMMARY: GERIATRIC PSYCHIATRY; EVALUATION STUDIES; DEPRESSION; ANXIETY; ALCOHOLISM

OBJECTIVES: The UPBEAT (Unified Psychogeriatric Biopsychosocial Evaluation and Treatment) program was established in 1994 to improve outcomes for older veterans hospitalized for medical conditions and who also had comorbid depression, anxiety, or substance abuse. The program used an intervention focused on treating and managing the veteran's concurrent mental disorder.

The UPBEAT program was initiated as a result of special Congressional

legislation mandating the VA program, and was specifically designed to end in 2001. As a result, the Coordinating Center ceased to exist at the end of FY01. At the request of VA Headquarters, SMITREC has agreed to maintain and utilize the UPBEAT data set. Additional IRB applications will be filed as required when specific plans for analyses and manuscript preparations are made.

RESEARCH DESIGN: Subjects were enrolled into the study at 9 VA Medical Centers nationwide. Patients were screened for depression, anxiety and substance abuse; those screening positive were randomized to intervention or control conditions. Follow-up assessments were conducted at 6-, 12-, and 24-month intervals following baseline.

METHODOLOGY: Data were coded and computerized at the West LA/UCLA Coordinating Center. All data cleaning and coordination has been the responsibility of the Coordinating Center. In addition, utilization data were extracted from the VA Administrative Data Sets at Austin and were merged with patient level data. Initial data analyses have been conducted, and several manuscripts have been published from the data set.

CLINICAL RELATIONSHIPS: Co-occurring physical and mental health disorders are a critical issue to ensure best practices in the treatment of older veterans in the VHA system. This database will provide VA headquarters, policy makers, and clinicians with needed techniques and information regarding the most effective means to deal with a serious issue in the care of this vulnerable older adult population. Per VA headquarters' request, SMITREC continues to maintain this database.

10/22/03 FINDINGS: UPBEAT intervention appears to accelerate the transition from inpatient to outpatient care (shorter lengths of stay) for patients admitted to acute medical or surgical hospital services who have undiagnosed psychiatric symptoms. The kind of case management or care coordination that appeared to be most successful in UPBEAT is similar to that done for other high cost patients by hospital-based home care managers and mental health intensive case management teams.

10/27/04 FINDINGS: UPBEAT intervention appears to accelerate the transition from inpatient to outpatient care (shorter lengths of stay) for patients admitted to acute medical or surgical hospital services who have undiagnosed psychiatric symptoms. The kind of case management or care coordination that appeared to be most successful in UPBEAT is similar to that done for other high cost patients by hospital-based home care managers and mental health intensive case management teams.

Two manuscripts have been submitted for publication and accepted, but have not yet been published: 1) Jarvik L, Gerson S, Maxwell A, Blow FC, et al. Symptoms of depression and anxiety (MHI) following acute medical/surgical hospitalization and post-discharge psychiatric (DSM) in 839 geriatric US veterans. *Int J of Geriatric Psychiatry* (in press); 2) Oslin DW, Thompson R, Kallan MJ, TenHave T, Blow FC, Bastani R, Gould RL, Maxwell AE, Jarvik L. Treatment effects from UPBEAT: A randomized trial of care management for behavioral health problems in hospitalized elderly patients. *J Geriatr Psychiatr Neurol* (in press).

DESCRIPTORS: GERIATRIC PSYCHIATRY; EVALUATION STUDIES; DEPRESSION; ANXIETY; ALCOHOLISM

18/5/17 (Item 8 from file: 266)

DIALOG(R) File 266: FEDRIP

Comp & dist by NTIS, Intl Copyright All Rights Res. All rts. reserv.

00499403

IDENTIFYING NO.: 158726; 0009; 583 AGENCY CODE: VA

Anxiety Symptoms in Schizophrenia: Association with Function, Self-Esteem and Symptoms

PRINCIPAL INVESTIGATOR: Lysaker, Paul, Ph.D.

PERFORMING ORG.: Department of Veterans Affairs, Medical Center, Indianapolis, IN

SPONSORING ORG.: Department of Veterans Affairs, Research and Development (15), 810 Vermont Ave. N.W., Washington, D.C. 20420 United States of America

DATES: 20040319

SUMMARY: SELF CONCEPT; SCHIZOPHRENIA; ASSOCIATION; ANXIETY

OBJECTIVES: The proposed study seeks to gather data on the co-occurrence of anxiety symptoms in schizophrenia with the primary long term goal of gathering pilot data to support a grant proposal for the development of a federally funded cognitive behavior therapy intervention

that might uniquely target anxiety symptoms in this group.

Initial OBJECTIVES:

1 To determine the frequency with which persons with schizophrenia experience significant levels of co-occurring anxiety symptoms including obsessions and compulsions, social anxiety and symptoms linked with trauma.

2 To determine whether co-occurring symptoms of anxiety are linked with deficits in working memory and executive function.

3 To determine whether co-occurring symptoms of anxiety are linked with lower levels of self-esteem, social function, and more avoidant coping as well as experiences of stigma.

4 To compare the severity of anxiety symptoms of persons with schizophrenia with those suffering from Post Traumatic Stress Disorder.

5 To determine the stability of anxiety symptoms and their correlates over a period of six months.

Long-term Objectives:

6 To gather pilot data on the need and possible targets for a Cognitive Behavior Therapy (CBT) intervention designed to reduce anxiety symptoms in schizophrenia.

7 To determine effect sizes necessary for a power analyses to determine sample sizes for a study of the effects of CBT on anxiety symptoms in schizophrenia. RESEARCH PLAN: Recruitment of 90 participants will begin upon approval. By enrolling a minimum of 35 participants per month we anticipate completing all initial assessments within four months. We expect to complete all follow-up assessments and all analyses within one year of receiving full approval for the protocol. We intend to apply for funding within one year of receiving full approval. We anticipate having completed and submitted all manuscripts from this study within 18 months of receiving study approval. METHODOLOGY: Ninety participants will be recruited from the psychiatry service of a Community Mental Health Center and VA Medical Center. To qualify for the study, participants must have a SCID confirmed diagnosis of schizophrenia or schizoaffective disorder and will be in a post acute stage of illness as defined by no hospitalizations, changes in type of psychotropic medication or in housing within the previous 30 days. Participants will be a minimum of 18 years of age. Exclusion criteria will include active substance use, or history of mental retardation diagnosis. For comparison purposes an additional 25 persons with Posttraumatic Stress Disorder will be recruited from the PTSD program of the VA Medical Center. Exclusion criteria for these participants will include a history of mental retardation or a diagnosis of a psychotic disorder or active substance abuse. To be eligible all PTSD participants must have had no hospitalizations, changes in type of psychotropic medication or in housing within the previous 30 days. Following informed consent and confirmation of eligibility, participant will undergo an initial assessment battery including measures of neurocognition, symptoms and function. Participants will be invited to return in 6 months for a reassessment, including all initial assessment procedures except neurocognitive assessment. In addition, at reassessment participants will be asked only about trauma experiences over the preceding six months.

Primary data analysis will include factor and cluster analyses to determine grouping of anxiety symptoms in the schizophrenia group. MANOVA and ANOVA procedures will be used to compare PTSD and schizophrenia participants on key measures. Multiple regression procedures will be utilized to examine links between level of anxiety symptoms and positive and negative symptoms, awareness of illness, neurocognition and psychosocial function. RESULTS: Most recently, with the use of federal funding

DESCRIPTORS: SELF CONCEPT; SCHIZOPHRENIA; ASSOCIATION; ANXIETY

18/5/18 (Item 9 from file: 266)

DIALOG(R)File 266:FEDRIP

Comp & dist by NTIS, Intl Copyright All Rights Res. All rights reserved.

00487767

IDENTIFYING NO.: 141878; 0016; 506 AGENCY CODE: VA

Maintenance of the UPBEAT data Set

PRINCIPAL INVESTIGATOR: Blow, Frederic C., Ph.D.

PERFORMING ORG.: Department of Veterans Affairs, Medical Center, Ann Arbor, MI

SPONSORING ORG.: Department of Veterans Affairs, Research and Development (15), 810 Vermont Ave. N.W., Washington, D.C. 20420 United States of America

DATES: 20010809

SUMMARY: SUBSTANCE-RELATED DISORDERS; MENTAL DISORDERS; ADULT

OBJECTIVES: The UPBEAT (Unified Psychogeriatric Biopsychosocial Evaluation and Treatment) program was established in 1994 to improve outcomes for older veterans hospitalized for medical conditions and who also had comorbid depression, anxiety, or substance abuse. The program used an intervention focused on treating and managing the veteran's concurrent mental disorder. RESEARCH PLAN AND METHODOLOGY: Subjects were enrolled into the study at 9 VA Medical Centers nationwide. Patients were screened for depression, anxiety and substance abuse; those screening positive were randomized to intervention or control conditions. Follow-up assessments were conducted at 6-, 12-, and 24-month intervals following baseline. 1377 subjects were randomized to the UPBEAT intervention and 1339 were included in the control group. Prior to enrollment, individual subjects signed informed consent. Data were coded and computerized at the West LA/UCLA coordinating Center. All data cleaning and coordination has been the responsibility of the Coordinating Center. In addition, utilization data were extracted from the VA Administrative Data Sets at Austin and were merged with patient level data. Initial data analyses have been conducted, and several manuscripts have been published from the data set. The UPBEAT program was initiated as a result of special congressional legislation mandating the VA program, and was specifically designed to end in 2001. As a result, the Coordinating Center ceased to exist at the end of FYO I. At the request of VA Headquarters; SMITREC has agreed to maintain and utilize the UPBEAT data set and will continue to do so indefinitely. Additional IRB applications will be filed as required when specific plans for analyses and manuscript preparations are made. CLINICAL RELEVANCE: Co - occurring physical and mental health disorders are a critical issue to ensure best practices in the treatment of older veterans in the VHA system. This database provides headquarters, policy makers, and clinicians with needed techniques and information regarding treatment interventions to deal with a serious issue in the care of this vulnerable older adult population. 08-12-2003 8/14/03 ls

Co - occurring physical and mental health disorders are a critical issue to ensure best practices in the treatment of older veterans in the VHA system. This database will provide VA headquarters, policy makers, and clinicians with needed techniques and information regarding the most effective means to deal with serious issue in the care of this vulnerable older adult population. Per VA headquarters' request, SMITREC continues to maintain this database.

Update 7/7/05: SMITREC continues to maintain the database at request of VA Central Office. There is no progress to report.

al
*** PDS Report: Progress; Report Date: 08/09/06; Submitted: 08/21/06
***There is no progress to report at this time although future analyses are planned.

al
*** PDS Report: Final; Report Date: 07/18/07; Submitted: 07/18/07 ***
FINAL REPORT 7/16/2007: No additional analyses have occurred in the past year, nor have there been any new publications from this data. The project is being terminated.

DESCRIPTORS: SUBSTANCE-RELATED DISORDERS; MENTAL DISORDERS; ADULT

18/5/19 (Item 1 from file: 56)
DIALOG(R)File 56:Computer and Information Systems Abstracts
(c) 2008 CSA. All rts. reserv.

0000030988 IP ACCESSION NO: 0041951
User-Responsive Subject Control in Bibliographic Retrieval Systems

Tague, J M
Univ. West. Ont., London, Can.

INFO. PROC. & MGMT., v 17, n 3, p 149-159, 1981
PUBLICATION DATE: 1981

DOCUMENT TYPE: Journal Article
RECORD TYPE: Abstract
LANGUAGE: English
FILE SEGMENT: Computer & Information Systems Abstracts

ABSTRACT:

A study was carried out of the relationship between the vocabulary of user queries and the vocabulary of documents relevant to the queries, and the value of adding to the document description record in a retrieval system keywords from previous queries for which the document had proved useful. Two test databases incorporating user query keywords were implemented at the School of Library and Information Science, University of Western Ontario. Clustering of the documents via title and user keywords, a statistical analysis of title-user keyword co-occurrences, and retrieval tests were used to examine the effect of the added keywords. Results showed the impracticality of the procedure in an operational setting, but indicated the value of analyses with sample data in the development and maintenance of keyword dictionaries and thesauri.

DESCRIPTORS: Bibliographic retrieval; Data base ; Clustering ; Keyword information

19/5/1 (Item 1 from file: 35)
DIALOG(R)File 35:Dissertation Abs Online
(c) 2008 ProQuest Info&Learning. All rts. reserv.

687919 ORDER NO: AAD80-16108

DESIGN OF A COMPUTER-BASED SYSTEM FOR RESEARCH IN AND TEACHING OF LITERATURE

Author: MADRON, BEVERLY BROWN
Degree: PH.D.
Year: 1979
Corporate Source/Institution: GEORGE PEABODY COLLEGE FOR TEACHERS (0074)
Source: VOLUME 41/01-A OF DISSERTATION ABSTRACTS INTERNATIONAL.
PAGE 234. 345 PAGES
Descriptors: LITERATURE, GENERAL
Descriptor Codes: 0401

Purpose. This paper describes the design, development, and partial implementation of the TUIT (Technology Utilized for Investigation and Teaching) system, a computer-based system for research and teaching in the field of literature.

The TUIT system at the present time consists of a basic calling and parameter-analyzing main program which links together four separate subprograms. The subprograms are designed to: modify and/or create files in a standardized format; index the file, either through KWIC or KWOC procedures, and provide an indication of the frequency of **appearance of words** in the text; **count the number of words, sentences, and syllables** in the text; and calculate several different readability indices for the textual material. Several additional programs and changes to existing programs are projected for the future development of the system.

TUIT, and other computer-based text handling systems, can be of value not only to the researcher, but also to the instructor, through the ability to provide consistent and comparable data about a single text or group of texts. The advantages to be gained--in terms of ease of handling the material, consistency of computations, and speed and accuracy in manipulating the text--make the required initial effort to convert the material to machine-readable form well worthwhile.

Appendices describing sample output from the system and providing both technical and non-technical guides to the use of the system are also included.

19/5/2 (Item 1 from file: 6)
DIALOG(R)File 6:NTIS
(c) 2008 NTIS, Intl Cpyrght All Rights Res. All rts. reserv.

0961454 NTIS Accession Number: AD-A113 962/5/XAB

A Computer Program for Assessing Readability

(Final rept)

Katznelson, J.

Human Engineering Lab., Aberdeen Proving Ground, MD.

Corp. Source Codes: 054879000; 172850

Report No.: HEL-TM-4-80

Feb 80 36p

Languages: English

Journal Announcement: GRAI8217

Order this product from NTIS by: phone at 1-800-553-NTIS (U.S. customers); (703)605-6000 (other countries); fax at (703)321-8547; and email at orders@ntis.fedworld.gov. NTIS is located at 5285 Port Royal Road, Springfield, VA, 22161, USA.

NTIS Prices: PC A03/MF A01

Country of Publication: United States

A computerized program for assessing the readability of technical documentation is presented. This program is particularly useful to Army personnel responsible for the readability of Army publications. The program is designed to provide the user with an analysis of the text that includes: (a) the complete text, (b) a listing of words containing 3 or more syllables and the number of times each multi-syllable word appears in the text, (c) the **number of sentences**, (d) the average sentence length, (e) the number of words, (f) the number of syllables, (g) the average syllables per word and (h) the Flesch-Kincaid reading grade level score. An appendix provides the reader with both a complete program listing (BASIC)

and sample input and output files.

Descriptors: *Reading; *Technical writing; *Human factors engineering; Documents; Military publications; Computer programs; Comprehension; Writing; Literacy; Automatic; Scoring; Army personnel; word lists; Output; Input; Files(Records); Quantity; Syllables; Length

Identifiers: Readability; Flesch reading ease formula; NTISDODXA

Section Headings: 88GE (Library and Information Sciences--General)

13/3,K/1 (Item 1 from file: 275)
DIALOG(R)File 275:Gale Group Computer DB(TM)
(c) 2008 The Gale Group. All rts. reserv.

01929334 SUPPLIER NUMBER: 18226838 (USE FORMAT 7 OR 9 FOR FULL TEXT)
Searching far and wide: the powerful document retrieval software of PLS.
(Personal Library Software) (Company Business and Marketing)
Banet, Bernard
Seybold Report on Desktop Publishing, v10, n8, p17(6)
April 22, 1996
ISSN: 0889-9762 LANGUAGE: English RECORD TYPE: Fulltext
WORD COUNT: 4746 LINE COUNT: 00379

... of a relevance score for each document are variables such as:
* The number of times each query term is found in a document;
* The number of different search terms that appear in a document;
* How close the 'hit' terms found are to the beginning of each document ;
* How closely together different search terms appear in the text;
* How closely the order in...

...the terms matching those in the query, as indicated by frequency of appearance in the document collection . (Rarer terms are more useful in indicating what a document is about, and receive a...

13/3,K/2 (Item 2 from file: 275)
DIALOG(R)File 275:Gale Group Computer DB(TM)
(c) 2008 The Gale Group. All rts. reserv.

01832377 SUPPLIER NUMBER: 17378357 (USE FORMAT 7 OR 9 FOR FULL TEXT)
What to do with documents.(Speed Reading: Optical Character Recognition Software)(document-management software) (Software Review)(Evaluation)
Computer Shopper, v15, n10, p529(1)
Oct, 1995
DOCUMENT TYPE: Evaluation ISSN: 0886-0556 LANGUAGE: English
RECORD TYPE: Fulltext; Abstract
WORD COUNT: 669 LINE COUNT: 00056

... more power than document managers for finding information in word processing, spreadsheet, and database files.
Text -retrieval programs are easier to set up and use because they handle fewer types of information and don't interact with...

...can start scanning and OCR from within document-management programs, you handle those tasks before starting a text -retrieval package, saving pages as text files.

The text-retrieval software then indexes those files, noting all the words it finds and the documents in which they appear. If you then search for, say, documents containing...

13/3,K/3 (Item 3 from file: 275)
DIALOG(R)File 275:Gale Group Computer DB(TM)
(c) 2008 The Gale Group. All rts. reserv.

01759224 SUPPLIER NUMBER: 16689579 (USE FORMAT 7 OR 9 FOR FULL TEXT)
Iota's Newswire: revolutionary archive and retrieval system. (Iota Industries Ltd Newswire text retrieval software)(includes overview on Iota)
Tribute, Andrew
Seybold Report on Publishing Systems, v24, n11, p3(11)
Feb 13, 1995
ISSN: 0736-7260 LANGUAGE: ENGLISH RECORD TYPE: FULLTEXT
WORD COUNT: 8463 LINE COUNT: 00639

... the letters appeared in reverse type.
Setting up a search. Above: This is how to set up a search for all articles containing the keyword "aids." Right: This page was found to contain the keyword "aids." This time, Newswire found the keyword in the headline, but...

...This color image is linked to a monochrome image created at the time of the initial page scan. In the initial implementation of Newsware, which is currently being delivered, color images are...

13/3,K/4 (Item 4 from file: 275)
DIALOG(R)File 275:Gale Group Computer DB(TM)
(c) 2008 The Gale Group. All rts. reserv.

01305363 SUPPLIER NUMBER: 07606922 (USE FORMAT 7 OR 9 FOR FULL TEXT)
Grammar checkers apply polish to writing's rough edges. (Software Review)
(Rightsoft Inc's Rightwriter, Lifetree Software Inc's Correct Grammar and Reference Software Inc's Grammatik III 1.1 grammar checkers; includes related article on RightWriter's lack of an interactive user interface) (evaluation)

Enyart, Bob; Erickson, Michelle; Webster, Steven; Frentzen, Jeffrey
PC Week, v6, n35, p35(3)
Sept 4, 1989
DOCUMENT TYPE: evaluation ISSN: 0740-1604 LANGUAGE: ENGLISH
RECORD TYPE: FULLTEXT; ABSTRACT
WORD COUNT: 2265 LINE COUNT: 00186

... detected and various readability indexes, including the education-level rating.

It also delivers the average number of sentences per paragraph, words per sentence and syllables per word. It presents the number of words, passive verbs, prepositions, question marks and exclamation points used in the document.

Grammatik III is...

13/3,K/5 (Item 1 from file: 636)
DIALOG(R)File 636:Gale Group Newsletter DB(TM)
(c) 2008 The Gale Group. All rts. reserv.

01896595 Supplier Number: 43295638 (USE FORMAT 7 FOR FULLTEXT)
BUREAU DATA BECOMING KEY TO RISK PREDICTION
Credit Risk Management Report, v2, n19, pN/A
Sept 14, 1992
Language: English Record Type: Fulltext
Document Type: Newsletter; Trade
Word Count: 999

... Services. TRW builds credit risk models using bureau information.

Behavior Models Formerly Looked Just at Master File Information
In behavior modeling, what typically occurred in the past was the use of master file information for account monitoring activities--such as credit line increases or decreases, authorizations and collections.

Behavior modeling has been effective for delinquencies of 30 and 60 days using only master file information, Esquinas said. But farther along, at 90, 120 or 180 days past due, looking at master file data was of limited value.

Collectible and non-collectible accounts both appear the same in terms of various characteristics from the master file-- balance, how long past due, how long someone...

...everybody begins to look homogeneous," Esquinas said.

The reliance on bureau information, in addition to master file data, was due to the frustration of collection agencies, financial institutions and retailers. The collection environment traditionally has offered a shotgun approach, Esquinas emphasized. At the point when the collection agency receives information from the financial institution, the account has been charged off.

Distinguishing accounts as to collectibility...

13/3,K/6 (Item 1 from file: 148)
DIALOG(R)File 148:Gale Group Trade & Industry DB
(c)2008 The Gale Group. All rts. reserv.

0017723499 SUPPLIER NUMBER: 126198603 (USE FORMAT 7 OR 9 FOR FULL TEXT)

Communication in Latin America: an analysis of Guatemalan business letters.

Conaway, Roger N.; Wardrope, William J.

Business Communication Quarterly, 67, 4, 465(10)

Dec, 2004

ISSN: 1080-5699

LANGUAGE: English

RECORD TYPE: Fulltext

WORD COUNT: 3769

LINE COUNT: 00339

... researchers independently translated the letters and agreed on the substantive content when translation of difficult words appeared. After translations to English were completed, each researcher independently counted the number of sentences in each letter sequentially and identified the sentence he thought stated the writer's central...

13/3,K/7 (Item 2 from file: 148)

DIALOG(R)File 148:Gale Group Trade & Industry DB

(c)2008 The Gale Group. All rts. reserv.

09841124 SUPPLIER NUMBER: 19781070 (USE FORMAT 7 OR 9 FOR FULL TEXT)

A behavioral model of nondurable consumption expenditure.

Tolar, Martin Michael

Journal of Socio-Economics, v26, n2, p291(12)

May-June, 1997

ISSN: 1053-5357

LANGUAGE: English

RECORD TYPE: Fulltext; Abstract

WORD COUNT: 4435

LINE COUNT: 00371

... interest rates and supposes that individuals are capable of making distinctions between economic variables in both nominal and real terms. Secondly real interest rates appear to be highly variable in the Australian experience over time, (ILLUSTRATION FOR FIGURE 2 OMITTED...

...and the methodologies employed in past examinations call into question the REPIH.

DATA AND METHODOLOGY

Data

Our investigation involves the collection and analysis of both primary and secondary data.

Primary Data

We employ the use of survey data which is obtained from the administration of a...

13/3,K/8 (Item 3 from file: 148)

DIALOG(R)File 148:Gale Group Trade & Industry DB

(c)2008 The Gale Group. All rts. reserv.

07200891 SUPPLIER NUMBER: 15232420 (USE FORMAT 7 OR 9 FOR FULL TEXT)

Document management pushing the paper aside.

Miller, Marlon

CMA - the Management Accounting Magazine, v68, n1, p13(3)

Feb, 1994

ISSN: 0831-3881

LANGUAGE: ENGLISH

RECORD TYPE: FULLTEXT; ABSTRACT

WORD COUNT: 2543

LINE COUNT: 00214

... include:

* Providing a central system function to organize, maintain and control access to the data base of electronic documents. For example, access to personnel records needs to be tightly controlled while corporate marketing strategies...

...distribution services, versioning services, etc.

* Providing several methods to intelligently access information -- full text retrieval (find all documents that contain these words or phrases), keyword indexing (find all ...for document management systems is growing at more than 35 per cent a year. Gartner Group estimates the worldwide market for document management software and integration will grow (excludes hardware) to \$2 billion by 1997 as the...

13/3,K/9 (Item 1 from file: 15)

DIALOG(R)File 15:ABI/Inform(R)

(c) 2008 ProQuest Info&Learning. All rts. reserv.

03243322 1117958741

Differential patterns of textual characteristics and company performance in the chairman's statement

Clatworthy, Mark A; Jones, Michael John

Accounting, Auditing & Accountability Journal v19n4 PP: 493 2006

ISSN: 0951-3574 JRNL CODE: AAJ

WORD COUNT: 8671

...TEXT: on equity (ROE). When analysing syntactical characteristics such as word count, syllables per word and **words per sentence**, they found that only word **count** statistically significant; high ROE firms were more verbose than low ROE firms. However, it is...

13/3,K/10 (Item 2 from file: 15)

DIALOG(R)File 15:ABI/Inform(R)

(c) 2008 ProQuest Info&Learning. All rts. reserv.

02995598 883131321

Elicitation and use of relevance feedback information

Vechtomova, Olga; Karamuftuoglu, Murat

Information Processing & Management v42n1 PP: 191-206 Jan 2006

ISSN: 0306-4573 JRNL CODE: IPM

...ABSTRACT: track of TREC-12. The first method consists of asking the user to select a **number** of **sentences** that represent documents. The second method consists of showing to the user a list of noun phrases extracted from the **initial document set**. Both methods then expand the query based on the user feedback. The TREC results show...

13/3,K/11 (Item 3 from file: 15)

DIALOG(R)File 15:ABI/Inform(R)

(c) 2008 ProQuest Info&Learning. All rts. reserv.

02981511 927286821

DETERRENCE VERSUS BRUTALIZATION: CAPITAL PUNISHMENT'S DIFFERING IMPACTS AMONG STATES

Shepherd, Joanna M

Michigan Law Review v104n2 PP: 203-255 Nov 2005

ISSN: 0026-2234 JRNL CODE: MLW

WORD COUNT: 18137

...TEXT: studies of capital punishment's deterrent effect and other capital punishment researchers have used similar **data sets**. I am now, for the first time, using the data to estimate separate deterrent effects for individual states.

we should expect some differences among the **data sets** in the states that fall into each group: deterrent, no effect, and brutalization. In some ...

...statistically significant effect that state-level data did not. The varying time periods of the **data sets** may also result in differences if states experienced deterrence or brutalization during some years, but not others. Nevertheless, the results from the other **data sets** can support the **primary data set**'s evidence that capital punishment has different impacts in different states.

A. State-Level Monthly...

...and other variables at the state level over the period 1977-1999. I used this **data set** in another recently published study in The Journal of Legal Studies. Because the data and...

...death row sentence in a given month is defined as a moving average of the **number** of death row **sentences** in the current and previous eleven months divided by a similar twelve-month moving average...

13/3,K/12 (Item 4 from file: 15)

DIALOG(R)File 15:ABI/Inform(R)
(c) 2008 ProQuest Info&Learning. All rts. reserv.

02851309 700532441

How effectively do marketing journals transfer useful learning from scholars to practitioners?

Crosier, Keith

Marketing Intelligence & Planning v22n5 PP: 540-556 2004

ISSN: 0263-4503 JRNL CODE: MIP

WORD COUNT: 6887

...TEXT: remainder of the document?" Click the No button.

A new dialogue box, Readability Statistics, will **appear** automatically.

It provides: **count of words**, characters, paragraphs and **sentences**; averages of sentences per paragraph, words per sentence and characters per word; a percentage figure...

13/3,K/13 (Item 5 from file: 15)

DIALOG(R)File 15:ABI/Inform(R)

(c) 2008 ProQuest Info&Learning. All rts. reserv.

02537032 268883891

An evaluation of help mechanisms in natural language information retrieval systems

Kreymer, Oleg

Online Information Review v26n1 PP: 30-39 2002

ISSN: 1468-4527 JRNL CODE: ONCD

WORD COUNT: 5578

...TEXT: Sometimes a system, claiming to use, for instance, "syntactic parsing" (i.e. analysis of the **sentence structure**), would end up **counting the appearance** of query **terms** in a document. As a result, retrieved documents were irrelevant to a search. This would...

13/3,K/14 (Item 6 from file: 15)

DIALOG(R)File 15:ABI/Inform(R)

(c) 2008 ProQuest Info&Learning. All rts. reserv.

02374531 124761131

Cutting costs or raising revenue?

Hollman, Lee

Call Center Magazine v15n5 PP: 50-62 May 2002

ISSN: 1064-5543 JRNL CODE: CCMA

WORD COUNT: 4360

...TEXT: employees.

Starting at \$10,000, the software's price varies with the number of **knowledge base articles** you create. The starting price entitles you to 500 articles. For approximately \$25,000, you...

...frames and JavaScript complicate keyword searches. So the spider includes natural language processing capabilities to **find** appropriate content from **each site**.

Customers can enter **keywords** and full sentences, and One Step can recognize typos without requiring you to save a...

13/3,K/15 (Item 7 from file: 15)

DIALOG(R)File 15:ABI/Inform(R)

(c) 2008 ProQuest Info&Learning. All rts. reserv.

02228661 81564673

Managed in Hong Kong: Adaptive Systems, Entrepreneurship and Human Resources

Selmer, Jan

ASEAN Economic Bulletin v18n2 PP: 247-249 Aug 2001

ISSN: 0217-4472 JRNL CODE: IAEB
WORD COUNT: 1116

...TEXT: Western connotations. Based on original empirical research, the author compares the extent of organizational commitment, both in behavioural and attitudinal terms. As expected, she finds that the Chinese exhibited a lower level on both counts than their western counterparts, but...

...stores, Yaohan, has closed its doors for the last time. Nevertheless, the chapter provides interesting primary data collected through in-depth interviews of both Japanese expatriates and local employees of Yaohan and Jusco...

13/3,K/16 (Item 8 from file: 15)
DIALOG(R)File 15:ABI/Inform(R)
(c) 2008 ProQuest Info&Learning. All rts. reserv.

01172548 98-21943
Supplement to 1995 ASIS annual meeting proceedings
Anonymous
American Society for Information Science. Bulletin v22n2 PP: 21-27 Dec 1995/Jan 1996
ISSN: 0095-4403 JRNL CODE: BAS
WORD COUNT: 5098

...TEXT: checking; search trees, result ranking and best match searching; and links to thesauri and related word strings generated by co-occurrence rankings.

Ray R. Larson, School of Library and Information Studies, University of California at Berkeley...overload. The Cheshire II system includes the following design features: * it supports SGML as the primary data base format of the underlying search engine; it is a client/server application where the...

13/3,K/17 (Item 9 from file: 15)
DIALOG(R)File 15:ABI/Inform(R)
(c) 2008 ProQuest Info&Learning. All rts. reserv.

01121155 97-70549
Cell-site design & EMF
Smith, Clinton
Cellular Business v12n12 PP: 52-54 Nov 1995
ISSN: 0741-6520 JRNL CODE: CLB
WORD COUNT: 1224

...TEXT: pro-active philosophy, however, is often more productive. This means setting up meetings with concerned groups and sending out preliminary information on the plans of the site and the benefits of cellular service to the public...

...consultants. Because it is a meeting of record and many concerned residents also may be present, every word that is spoken by the company representative will be analyzed and interpreted. Having standard rehearsed...

16/3,K/1 (Item 1 from file: 16)
DIALOG(R)File 16:Gale Group PROMT(R)
(c) 2008 The Gale Group. All rts. reserv.

13909820 Supplier Number: 160640487 (USE FORMAT 7 FOR FULLTEXT)
Pathway research in plants.(Instruments & Systems: Bioinformatics Focus)
Bioscience Technology, v32, n2, p24(1)
Feb, 2007
Language: English Record Type: Fulltext
Document Type: Magazine/Journal; Trade
Word Count: 105

(USE FORMAT 7 FOR FULLTEXT)
TEXT:
...000 plant-specific abstracts and four full-text plant research journals.
In addition, 382,000- **sentence co - occurrence** facts for plant proteins
are available.

16/3,K/2 (Item 1 from file: 148)
DIALOG(R)File 148:Gale Group Trade & Industry DB
(c)2008 The Gale Group. All rts. reserv.

07238482 SUPPLIER NUMBER: 14989940 (USE FORMAT 7 OR 9 FOR FULL TEXT)
Poststructuralism and the ARTFUL database: some theoretical considerations.
(Project for American and French Research on the Treasury of the French
Language)
wolff, Mark
Information Technology and Libraries, v13, n1, p 35(8)
March, 1994
ISSN: 0730-9295 LANGUAGE: ENGLISH RECORD TYPE: FULLTEXT; ABSTRACT
WORD COUNT: 4551 LINE COUNT: 00370

... de nous; mais aussi nous n'aimons, nous n'embrassons rien de
reel.[12]
This **sentence** was located in the ARTFL database with a **co -**
occurrence search for the patterns "idE[aeo],*", "rE[ea]l.*" and "quant."
I specifically wanted to...

16/3,K/3 (Item 1 from file: 15)
DIALOG(R)File 15:ABI/Inform(R)
(c) 2008 ProQuest Info&Learning. All rts. reserv.

03199776 1199058321
Effect of relationships between words on Japanese information retrieval
Matsumura, Atsushi; Takasu, Atsuhiro; Adachi, Jun
ACM Transactions on Asian Language Information Processing v5n3 PP: 264
Sep 2006
ISSN: 1530-0226 JRNL CODE: TLIP

...ABSTRACT: words in a sentence. The second method uses proximity
relationships, particularly information about the ordered **co - occurrence**
of words in a **sentence**, to approximate the dependency relationships
between them. A Structured Index has been constructed for these...

16/3,K/4 (Item 2 from file: 15)
DIALOG(R)File 15:ABI/Inform(R)
(c) 2008 ProQuest Info&Learning. All rts. reserv.

02631629 403873091
Using the Co-occurrence of words for Retrieval Weighting
Mittendorf, Elke; Mateev, Bojidar; Schauble, Peter
Information Retrieval v3n3 PP: 243-251 Oct 2000
ISSN: 1386-4564 JRNL CODE: NFRT

...ABSTRACT: of our experiments for weighted retrieval is the surprising
result that features that describe the **co - occurrences** of words in
sentence -size or paragraph-size windows are significantly better

descriptors than purely word-based indexing features...

16/3,K/5 (Item 3 from file: 15)
DIALOG(R)File 15:ABI/Inform(R)
(c) 2008 ProQuest Info&Learning. All rts. reserv.

00850849 95-00241

Poststructuralism and the ARTFL Database: Some theoretical considerations

Wolff, Mark

Information Technology & Libraries v13n1 PP: 35-42 Mar 1994

ISSN: 0730-9295 JRNL CODE: JLA

WORD COUNT: 4174

...TEXT: de nous; mais aussi nous n'aimons, nous n'embrassons rien de
reel.(12)

This sentence was located in the ARTFL database with a co - occurrence
search for the patterns "idE[aeo].*", "rE[ae]l.*" and "quant" I
specifically wanted to...